8-27-2009

On the sources of astrometric anomalous refraction

Malinda Suzanne Taylor

Follow this and additional works at: https://digitalrepository.unm.edu/phyc_etds
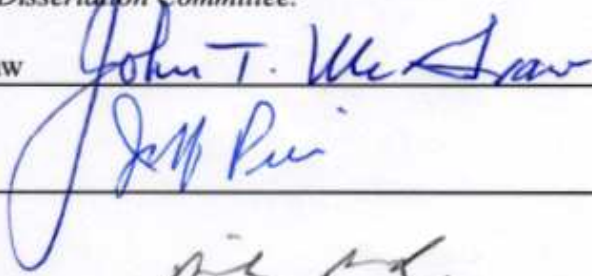
M. Suzanne Taylor

Candidate

PHYSICS AND ASTRONOMY

Department

This dissertation is approved, and it is acceptable in quality
and form for publication:
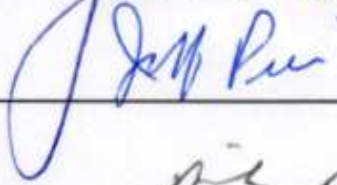
Approved by the Dissertation Committee:

Dr. John McGraw   , Chairperson

Dr. Jeff Pier

Dr. Rich Rand

Dr. Dinesh Loomba

ON THE SOURCES OF ASTROMETRIC ANOMALOUS REFRACTION

BY

M. SUZANNE TAYLOR

B.S., Physics, Linfield College, 2002
M.S., Physics, University of New Mexico, 2005

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Physics

The University of New Mexico
Albuquerque, New Mexico

August, 2009

DEDICATION

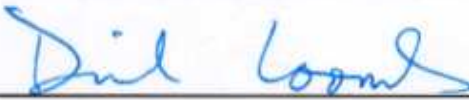This dissertation is dedicated to all who have supported me throughout my long educational career. I would especially like to thank my Mom who has encouraged me to follow my dreams and been proud of me whatever I did, and my Dad who is a constant source of inspiration and taught me to keep a sense of humor no matter what life throws at you. To my brother, Scott, thanks for the motivation – you may make more money than I do, but I'm smarter and more educated. Finally, I owe a great deal of gratitude to the New Mexico Touring Society for helping me maintain both my mental and physical health during my graduate career. To Johanna, you've been like a second mother to me and our Sunday morning chats on the trail have been a highlight of my time here. And to the rest of NMTS, Sandy, Susan, Ed, Will, Bob, Fred and everyone else, I'll miss riding with you – it's been fun.

ACKNOWLEDGMENTS

ON THE SOURCES OF ASTROMETRIC ANOMALOUS REFRACTION

BY

M. SUZANNE TAYLOR

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Physics

The University of New Mexico
Albuquerque, New Mexico

August, 2009

ON THE SOURCES OF ASTROMETRIC ANOMALOUS REFRACTION

by

M. Suzanne Taylor

B.S., PHYSICS, LINFIELD COLLEGE, 2002
M.S., PHYSICS, UNIVERSITY OF NEW MEXICO, 2005
PH.D., PHYSICS, UNIVERSITY OF NEW MEXICO, 2009

ABSTRACT

Over a century ago, astronomers using transit telescopes to determine precise stellar positions were hampered by an unexplained periodic shifting of the stars they were observing. With the advent of CCD transit telescopes in the past three decades, this unexplained motion, now known as "anomalous refraction," is again being observed.

Anomalous refraction is described as a low frequency, large angular scale motion of the entire image plane with respect to the celestial coordinate system as observed and defined by previous astrometric catalogs. These motions of typically several tenths of an arcsecond with timescales on the order of ten minutes are ubiquitous to drift-scan ground-based astrometric measurements regardless of location or telescopes used and have been attributed to the effect of tilting of equal-density layers of the atmosphere. The cause of this tilting has often been attributed to atmospheric gravity waves, but never confirmed. Although theoretical models of atmospheric refraction show that atmospheric gravity waves are a plausible cause of anomalous refraction, an observational campaign

specifically directed at defining this relationship provides clear evidence that anomalous refraction is not consistent with the passage of atmospheric gravity waves. The source of anomalous refraction is found to be meter scale slowly evolving coherent dynamical structures in the boundary-layer below 60 meters.

Table of Contents

List of Figures

1. Introduction

Astrometry is the key to understanding our place in the universe. It can truly be considered the original astronomy, existing long before telescopes and CCDs, dating back to a time when the universe was defined by myths and legends and the Earth was the center of a very small cosmos. In an age when the true nature of the stars, planets and galaxies was completely unknown, astronomers measured the skies with incredible precision using only their eyes. The Polynesians crossed the ocean using only their expansive knowledge of stellar positions to navigate. At the same time the Ancient Pueblo People at Chaco Canyon in the North American Southwest applied their understanding of celestial motions to precisely align their architecture with the cardinal directions and the locations of sunrise and sunset on the solstices. The start of the astronomical renaissance was ushered in by Tycho Brahe who made extensive systematic observations of stellar and planetary positions with unprecedented accuracy (less than one arcminute using the unaided eye).

The field of astrometry has been vital throughout history and throughout the world for navigating and charting the globe, marking the seasons and defining calendars. Mythologies of nearly every culture are based on an intimate knowledge of the positions of the stars and any changes in the heavens were once immediately recognized and heralded as harbingers of good or evil, omens of things to come. Only in the modern era has the public perception of the night sky dimmed to a mere recognition of its existence,

while the knowledge of the positions and motions of its denizens remains only in the hands of a miniscule fraction of the general population.

Ironically, at the current height of the scientific era with the explosion of telescope technology and the exponential increase in our understanding of the nature of the universe, it has become more vital than ever before that we know to incredible precision not only the positions of celestial bodies but their subtle parallaxes and motions as well. Modern astrometry is much more than just celestial map-making, although knowing the positions of stars, galaxies, *etc.* for the purpose of telescope pointing is not to be discounted. High precision positional measurements have shown that every object in the universe moves and those motions can reveal much about the nature of the universe. By measuring the motions of nearby stars we learn not only about the galaxy, but also about solar origins (*e.g.* Magnier *et al.* 2008). The measurement of parallaxes provides us the key to unlocking the scales of the universe (*e.g.* Herschel & Banks 1782). Observing the wobble of stars with planetary systems could tell us the nature of the planets we can't observe directly, and has proven to be a new means of detecting these systems in the first place (*e.g.* Sozzetti 2005). Measuring the motions of an asteroid allows us to calculate the orbit and determine if it may one day impact Earth (*e.g.* Bowel 2005). This is just a subset of the vital contributions astrometry makes to the astronomical community.

In light of these many relevant pursuits, in recent decades the astronomical community has devoted itself to building new major infrastructure specifically for the advancement of astrometry. Space-based missions such as the Space Interferometry Mission Lite

(SIM-Lite, *e.g.* Goullioud, *et al.* 2009) and terrestrial instruments like the Large Synoptic Survey Telescope (LSST, Tyson 2002) have the potential to launch a new era in our understanding of the universe.

There is, unfortunately, a long standing mindset in the field of astrometry that all of the errors inherent in astrometric observations are understood or insignificant. But upon close inspection of these "understood errors" reveals a gaping hole in the list – the atmosphere. This is not to say that astrometrists are ignorant about the effect of the atmosphere on ground-based optical observations - there is simply a long standing misconception that the atmosphere is not a source of positional errors in astrometric observations. Most astrometrists assume that any positional errors due to the atmosphere will average out over the course of an observation or several observations. Observations made in the last decade and even dating back 100 years suggest quite the contrary, the atmosphere may actually be the leading cause of systematic positional errors in astrometric measurements, and that these errors may result from waves or other dynamics of the lower atmosphere.

In the field of atmospheric physics, atmospheric gravity waves (AGWs) have long been a central subject in the study and understanding of atmospheric dynamics. In the field of astrophysics on the other hand, the atmosphere is well known as the prime limiting factor in ground based imaging capabilities. While astronomers are familiar with gravity waves (g-modes) in stellar atmospheres, atmospheric gravity waves and their major dynamical influence on the Earth's atmosphere are virtually unknown. These waves and their

possible refractive modulations of astronomical images have been acknowledged in only a small handful of papers spread over more than a century. Despite their astronomical obscurity, atmospheric gravity waves may be one of the largest consistent sources of error in ground-based astrometric measurements.

1.1 Historical background

At the end of the 19th century a great deal of attention was focused on geodesy and the definition of an astronomical coordinate system that accounts for the highly mobile platform of our terrestrial reference frame. In particular, the independent discoveries by Chandler (1885) and Küstner (1888) of a systematic variation in latitude caused by quasi-annual drift of the Earth's axis of rotation by several tens of meters (polar motion) led to an international observing program to measure these changes in latitude at locations across the globe (Schlesinger 1899). Latitude measurements were made using transit instruments such as the visual zenith telescope or the almucantar (Chandler 1887) wherein the measured transit time across a small circle centered on the zenith of a star of known celestial position gives the latitude of the observer's position.

Because the effect of normal zenith-angle refraction due to the plane-parallel atmosphere was well known at this time, additional concern was being voiced regarding the potential for errors in stellar position measurements due to tilting of these equal-density layers (*e.g.* Helmert and Albrecht 1898, Schlesinger 1899). For several years following the polar motion discoveries by Chandler and Küstner, many astronomers actually expressed concern that the yearly variations in latitude were actually caused by annual variability in refraction (Schlesinger 1916). Schlesinger actively addressed this issue of "anomalous

refraction" in his 1905 paper specifically devoted to examining systematic errors in measuring the variation of latitude. He suggested that simultaneous measurements of latitude by two observers on radically different instruments at the same site would be an ideal test of whether the source of anomalous refraction (AR) was actually external to the instrument-observer system. Precisely such a measurement was made (although, not for the purposes of studying AR) several years prior and indicated that a common error was affecting both observations (Schlesinger 1905). At this time accounts of anomalous refraction (AR) were based on isolated observations wherein the altitudes of a set of stars were unusually high or low by a small fraction of an arcsecond. The nature of these observations prevented recognition of any more than an hypothesized seasonal temporal component of the supposed effect.

In 1906 Schlesinger continued his research with a theoretical examination of how tilted atmospheric strata cause anomalous refraction and compared his theory with results from the international latitude observing program. Perhaps assuming that AR is a constant or very long period effect, as described above, Schlesinger used means of errors from hundreds of observations spread over a two year period from each of the latitude observing stations to look for anomalous refraction (Schlesinger and Blair 1906, Schlesinger and Hudson 1916). His result that the errors attributed to AR are not more than several hundredths of an arcsecond and therefore inconsequential when compared with observer and telescope errors was perhaps to be expected.

Anomalous refraction was first considered as a time varying phenomenon on short timescales by Perrine (1913). In discussing the challenges imposed on visual parallax measurements by irregularities in atmospheric refraction with timescales of seconds or less, he noted that there may also be similar positional displacements occurring on timescales of several minutes to hours.

ELECTRA, PLATE 622

ALCYONE, PLATE 617

ELECTRA, PLATE 631

ALCYONE, PLATE 631

ELECTRA, PLATE 640

ALCYONE, PLATE 640

ALCYONE, PLATE 650

0   20   40   60   80   100   120   140   160 seconds.

IGURE I. Trails of the Pleiades on Yerkes Plates. Vertical scale, one inch = 1″.2, or one centimeter = 0″.46.
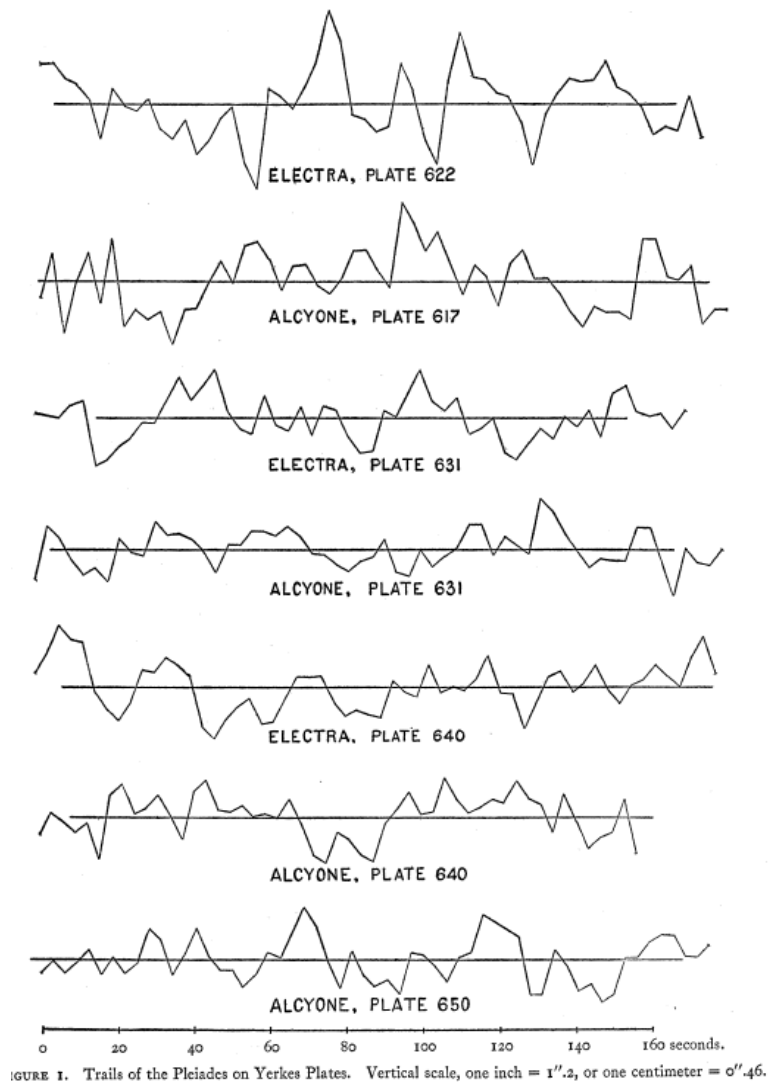
Figure 1. Positional fluctuations in star trails observed by Schlesinger (1916).

Analysis of AR occurring on minute timescales was pursued using photographic records of star trails by both Schlesinger (1916, Figure 1) and Hudson (1929). Schlesinger

6

acquired multiple trail plates of the Pleiades taken with the 40 inch refractor at the Yerkes observatory employing extensive precautions to ensure the telescope was completely stationary. As a comparison, he also obtained trail plates during which the telescope "was struck a sharp blow with the fist in the direction of declination" to examine any instrumental influences on the motions of the star trails. The star trails, although complicated by rapid seeing undulations, were found to exhibit both latitudinal and longitudinal oscillations of minute timescales (see Figure 1). The effect of striking the telescope was to induce oscillations with periods of less than a second which subsided rapidly. Hudson briefly continued his previous research with Schlesinger by conducting experiments to define the spatial scales of anomalous refraction. By observing star trails simultaneously with two cameras separated by 12 inches and then the same two cameras separated by 55 meters he showed that the anomalous refraction observed by the close cameras was highly correlated (Figure 2) while far less agreement was seen by the widely separated cameras (Figure 3).
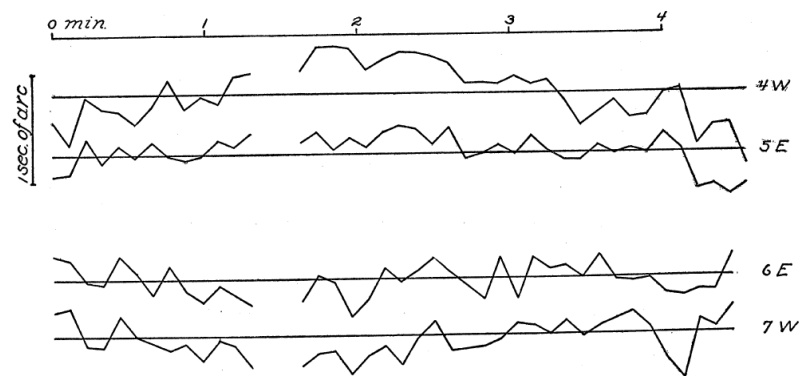
FIGURE I. Simultaneous trails of the same star with two cameras side by side.

Figure 2. Simultaneous star trails from two side-by-side cameras (Hudson, 1929)

Perhaps some of the most significant results to emerge from this early research were the experimental separation of AR from instrumental or observer errors and the subsequent

determination of limiting spatial scales for the effect. From these results, astronomers involved in the measurement of latitudes, parallaxes and positions recognized the importance of maintaining homogenous atmospheric conditions in the immediate vicinity of the telescope. The potential was acknowledged that some or all of anomalous refraction could be caused by atmospheric anisotropies or "room refraction" in or around the telescope building (Lambert, Schlesinger and Brown, 1931).
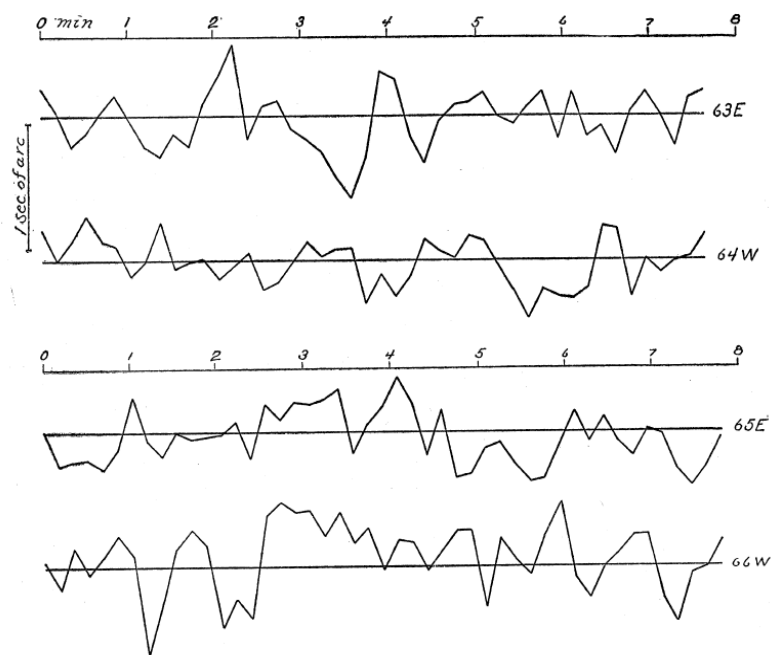
FIGURE 4. Simultaneous trails of the same star in two widely separated cameras.

Figure 3. Simultaneous star trails from two cameras separated by 55 meters (Hudson, 1929)

In the middle of the last century Land (1944, 1954) published two studies on the effects of anomalous refraction on photographic parallax measurements. In his first paper he discussed yearly and daily components of anomalous refraction as the potential causes of systematic errors between observations made on different nights. Land's second paper addressed the minute timescale component of anomalous refraction with experiments geared towards elucidating periods and amplitudes of the oscillations as well as

dependence on other sources of error such as measuring technique, telescope vibration and atmospheric turbulence. In particular, he noted no apparent variation in the observed anomalous refraction between nights with good seeing and those with poor seeing, suggesting that the two phenomena are unrelated.

The long term variations in refraction as observed in latitude measurements were again addressed by Sugawa in a series of papers. In his 1956 publication Sugawa used Radiosonde profiles to compute the varying tilt of atmospheric layers with altitudes up to 19 kilometers as a function of season and estimated how similar tilts may have affected previous latitude measurements. Sugawa's 1958 publication addressed the relationship between local wind speeds and directions and anomalous refraction, while his 1960 paper determined that the local Z term, or the difference between observed variation in latitude and that calculated from polar motion, is due to anomalous refraction.

Few significant references to anomalous refraction or any comparable atmospheric effects are to be found between 1960 and the 1990's. This lack of interest may be accounted for by replacement of traditional transit circles, meridian circles and other older astrometric methods with more modern photographic techniques. Observations of anomalous refraction become prevalent again in the last decade of the 20th century with all earlier work on the subject seemingly forgotten by the astronomical community.

The occurrence of anomalous refraction in the optical regime has been recently noted by several astrometric projects. During a study to determine stars positioned to be occulted

by Pluto, Dunham, McDonald and Elliot (1991) noted a "low-frequency motion of the sky coordinate system relative to the CCD" with a peak-to-peak amplitude of a few tenths of an arcsecond. They state that the motion could result from either local (telescope wind loading, mirror wind loading, etc.) or atmospheric causes. Stone *et al.* (1996) describe a quasi-periodic anomalous refraction in positions referenced to a nightly mean encountered during testing of the Flagstaff Astrometric Scanning Transit Telescope (FASTT), with periods ranging from a few minutes to several tens of minutes (Figure 4). The refraction was determined to be entirely atmospheric in nature, because not only is the FASTT telescope highly stable, but all telescope drives were inactive and the scale of the refraction increased with zenith distance. The Carlsberg Meridian Telescope (CMT) also observed similar refraction effects when undergoing a drift-scan survey (Evans, Irwin and Helmer 2002, Figure 5).
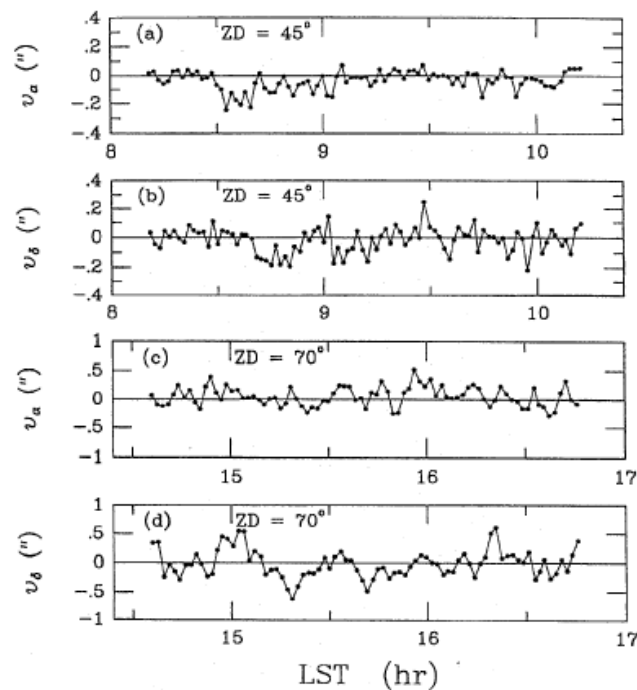
Figure 4. FASTT observed anomalous refraction at two different zenith angles in residuals referenced to a nightly mean.

10
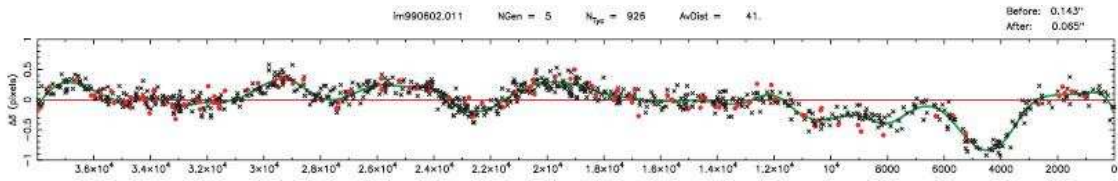
Figure 5. Anomalous refraction observed with the Carlsberg Meridian Telescope when referenced to Tycho-2 (Evans, Irwin and Helmer 2002).

Pier *et al.* (2003) included a comprehensive discussion of anomalous refraction observed during the commissioning of the Sloan Digital Sky Survey (SDSS) in his analysis of the astrometric calibration of that telescope. He noted quasi-periodic residuals when comparing stellar positions observed by the Sloan telescope with both the Tycho-2 and US Naval Observatory CCD Astrograph Catalog (UCAC) astrometric catalogs. Spline fits between each dataset and the catalog remove motions with timescales longer than about ten minutes, but minute timescale motions are clearly evident in the residuals (Figure 6). These residuals were described as having peak-to-peak amplitudes of tenths of an arcsecond and quasi-periods of a few to several tens of minutes. Comparison of residuals from each of the CCDs in the focal plane showed a high degree of consistency across the 2°.3 array. This would indicate that if the source of this motion is atmospheric in nature, it must be caused by atmospheric distortions which are coherent over at least this scale on the sky. Both Stone and Pier suggest that the source of the observed refractions may be atmospheric gravity waves occurring at altitudes of a few hundred meters up to 2 km.

11
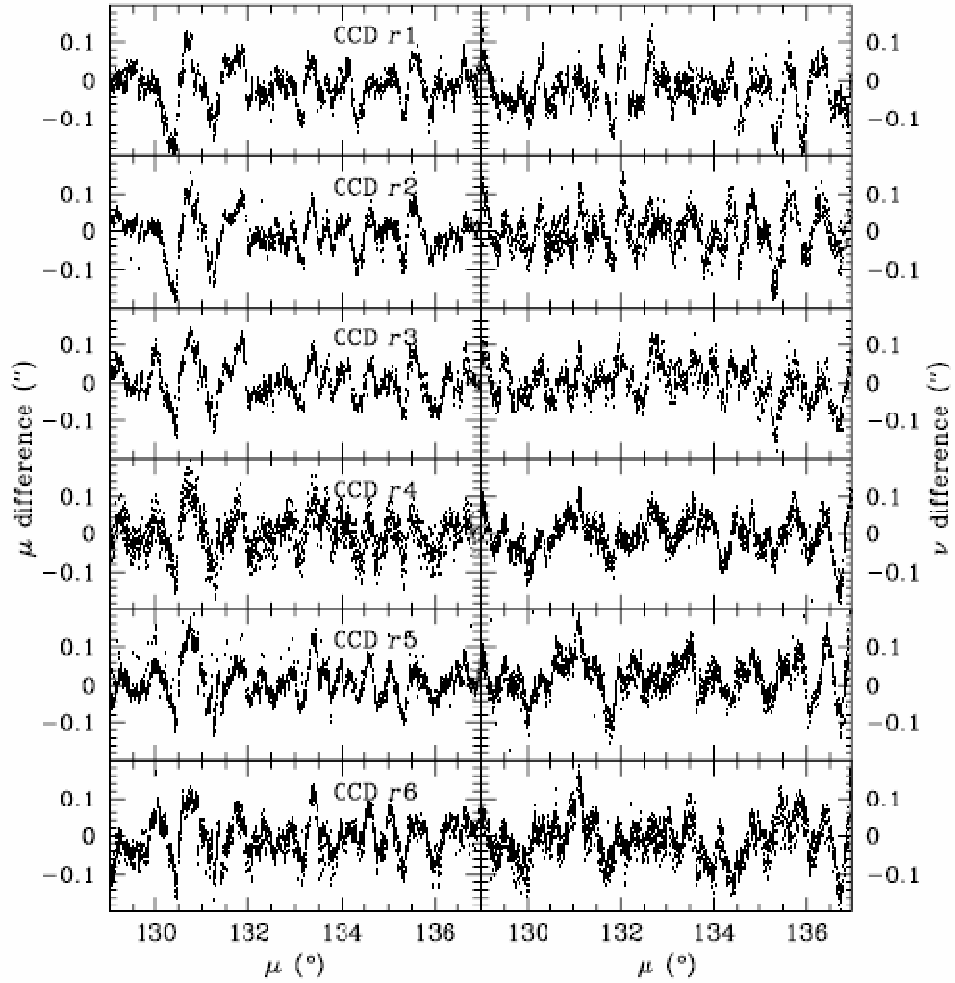
Figure 6. Anomalous refraction observed with the SDSS when referenced to Tycho-2 (Pier *et al.* 2003). μ and ν are approximately RA and Dec respectively and the differences are between star positions and catalog positions. Each row of images in the figure corresponds to the residuals in one of the *r*-band CCDs (the six *r* CCDs make up the first row of the SDSS 30 CCD focal plane array).

The most recent study geared exclusively towards the understanding of anomalous refraction was completed by Hirt (2006). Having encountered the anomalous refraction effect in his astrogeodetic observations, Hirt endeavored to improve the quantitative understanding of anomalous refraction at very low frequencies (~20 minutes to a few hour periods). Six nights of observations of several thousand stars using a digital zenith camera demonstrated anomalous refraction occurring consistently with the above periods

12

and with amplitudes of about 0.1 arcsecond. Hirt attributes the observed oscillations to the influence of the heterogeneous nature of the observing site on the atmosphere.

1.2 Motivation

The term "anomalous refraction" generically refers to any refraction which varies from that which is solely a function of the zenith angle. Due to the radial density gradient of the atmosphere, light entering the atmosphere at an angle will be refracted such that a star will appear higher in the sky than it actually is. This is considered normal atmospheric refraction. Although the effect of normal atmospheric refraction is significant, it has been extensively studied (*e.g.* Gubler & Tytler 1998) and a basic knowledge of atmospheric structure can be applied during the image reduction phase to correct for any non-anomalous position errors.

In a brief survey of modern ground based astrometry projects, five (including the four above mentioned above) definitely observed an anomalous refraction in their data (Pier *et al.* 2003; Evans *et al.* 2002; Stone *et al.* 1996; Guseva 1995; Dunham *et al.* 1991). One project (CTI, Benedict *et al.* 1991) may have observed anomalous refraction but requires further consideration, and two (UCAC, Zacharias *et al.* 2000 and Spacewatch, Gehrels *et al.* 1986) did not see anomalous refraction in their data. All five which observed the refraction were operating in Time-Delay and Integrate (TDI, also known as drift-scan) mode. Of the two projects which did not encounter any anomalous refraction, UCAC was operating in stare mode (Zacharias *et al.* 2000), while Spacewatch operated in TDI mode but suffered extremely poor internal consistency (uncertainty as high as 2") (Gehrels *et al.* 1986). The CCD/Transit Instrument (CTI) was operated in TDI mode, but

only used short (few arcminute) segments of the strip for the limited astrometry performed (Benedict *et al.* 1991). While no anomalous refraction was observed, further examination of the CTI data on larger angular scales may reveal this effect (see Chapter 3).

Rather than actively pointing the telescope at a celestial coordinate and exposing the CCD while tracking the diurnal sky motion ("stare mode"), a telescope operating in TDI mode remains fixed (or, in certain cases such as the SDSS, is tracked at a non-sidereal rate) during an observation while the sky tracks overhead. The CCD is clocked such that the charge is moved across the device at the apparent sidereal rate. This allows the telescope to observe continually as long as darkness and weather permit, with no dead time for the CCD to read out. TDI mode will be discussed in more detail in Chapter 3.

Older projects which observed anomalous refraction typically employed transit or meridian circles, images of star trails, or repeated photographic observations of the same field. Transit and meridian circles were typically operated with an observer recording the precise time at which a star was bisected by a vertical crosshair in the field of view of the telescope. A number of measurements would be taken of a single star during the star's transit of the meridian. This was accomplished by having either a number of vertical wires separated by a few arcminutes or a traveling wire which would remain centered on the star as it crossed the field of view and record the instants the wire passed through certain points on the field (Watts 1960). These older methods are comparable to modern

14

drift-scan operations in the sense that they inherently record information in both the temporal and angular domains.

Based on the time periods and scales over which the image motion due to anomalous refraction occurs, it seems likely that this is an effect which may *only* be observed by a telescope operating in TDI mode (or similar time-resolved operation). A telescope operating in stare mode would have to expose an image for at least as long as the period of the refraction to observe any image motion. If observed, the motion would appear simply as a smearing of the image or "guiding error". The image motion due to atmospheric turbulence will generally be of greater amplitude than anomalous refraction, and will overpower any smearing due to the latter (see Section 2.2.1 for more on atmospheric turbulence). A stare mode telescope making short exposure time images will only see the anomalous refraction if comparing multiple images from the same or similar altitudes and azimuths, with appropriate integration times and observing cadences to sample the effect in the time domain.

Additionally, many modern telescopes operating in stare mode are actively guided to remove any atmospheric or structural motions. Guiding can occur either by the simple expedient of adjusting the position of the telescope to maintain the position of a reference star relative to the field of view of the telescope, or through more technically intensive adaptive or active optics systems involving natural or laser guide stars. In the case of artificial guide star adaptive optics, a sodium laser beacon is directed along the line of sight of the telescope and excites sodium ions in the upper atmosphere, causing them to

glow, creating an artificial "star" (Thompson & Gardner 1987). Relative motions between program stars in the field of view and the artificial star are due to a combination of atmospheric "tip-tilt", seeing and telescope motion (if the laser is not a part of the telescope structure), and are removed in near real-time by the telescope optics and guiding. If records are maintained of any corrections made by the guiding system, it is likely that the anomalous refraction signature would be present. Indications from the Gemini North telescope are that these signatures have been seen by the laser guide star system (Laycock, private communication), but the source of the motions remains undetermined.

If a TDI telescope is pointed such that the time for an object to transit the field of view is on the order of one minute, individual objects may be displaced due to anomalous refraction but will not significantly move over the course of the exposure. Only over an extended observation of duration comparable to or longer than the wave period, during which one or more wavelengths of an AGW may pass over the telescope, is the anomalous refraction likely to be noted. The effect of the refraction will be a gradual, roughly periodic shifting of the images along the strip. An example of this might be a strip where the stars at a particular RA are shifted by 0.1" toward a higher RA while at an RA ten minutes later along the strip the stars are shifted by 0.1" toward a lower RA.

Median seeing induced by atmospheric turbulence is of the order one arcsecond at most observatories. Thus, a legitimate question might be why an effect that is often a factor of ten smaller than the seeing effects should be of concern. This is a perfectly reasonable

question, and in most cases of astronomical observations, where one is interested only in the clarity of an image or the signal-to-noise ratio of an object, anomalous refraction is of no consequence. However, when the purpose of study is to determine highly precise global positions of celestial objects, any induced shift of the image can undermine the effort. It is important to note that the primary effects attributed to seeing, *i.e.* scintillation and image motion, occur on timescales of fractions of a second and can be expected to average out over the length of the exposure. The result of seeing is a notable broadening of the point spread function (PSF) but no actual net displacement of the peak. The observed effect referred to as anomalous refraction has timescales of minutes or more and is systematic in nature; therefore it cannot be expected to average out, even on the longest (stare) exposures. What results is a systematic displacement of the apparent position of the peak of a star's PSF to a degree which will significantly reduce the accuracy and precision of an astrometric observation.

Having already been noted as a source of error in a number of astrometric surveys including the SDSS (Pier *et al.* 2003), the importance of a thorough investigation of anomalous refraction becomes increasingly evident when considering the upcoming and planned very large-scale imaging photometric surveys such as CTI-II (McGraw *et al.* 2005), PanSTARRS (Kaiser *et al.* 2002; Chandler 2004) and LSST (Tyson 2002), for which astrometry is fundamental. For these surveys, the value of the image data is dramatically enhanced by the ability to produce calibrated global astrometry to well below arcsecond precision and accuracy. Of greatest importance in this investigation, beyond just understanding the effect, is obtaining a comprehensive understanding of the

physical cause of anomalous refraction. An understanding of a problem allows for the treatment of the problem, but understanding the cause of the problem allows for its correction and prevention in the first place.

The research discussed in this thesis was designed to elucidate the true nature of anomalous refraction, first by experimentally answering the question of whether AR is caused by atmospheric gravity waves, and then by further clarifying the specific atmospheric dynamics and their associated optical properties responsible for AR. From this we hope to be able to determine the atmospheric conditions under which AR is most likely to occur and potentially facilitate a more effective approach to minimizing or reducing the effects of anomalous refraction. In Chapter 2 we discuss the physics of atmospheric refraction, atmospheric gravity waves and other atmospheric dynamics; particularly addressing how different atmospheric conditions may affect an astronomical image and characteristics we should look for in our data. Chapter 3 covers our analyses of pre-existing astronomical and atmospheric data to better understand the characteristics of AR, including rate of occurrence, as well as typical conditions in the nocturnal atmosphere. In Chapter 4 we describe the observations made for this research, both astronomical and atmospheric. Results will be given in Chapter 5 and Chapter 6 will cover a discussion of these results.

2. Conceptual Model

The first stage of this research is to gain a theoretical understanding of the atmosphere and how it serves as an optical element in every ground-based telescope system.   At the center of our study is a detrimental astrometric effect of unknown origin which we desire to understand.

One of the most effective approaches to problem solving is to work through the process in which the problem occurs step by step, examining every possible source of error until you have solved the problem. This approach can also be applied to problem-solving in most experimental or observational systems. For our particular case we need to consider the path that the light takes from a star at the zenith, starting at the top of the atmosphere and ending with the position derived from its image. This path includes all parts of the atmosphere, the observatory, the various elements of the telescope, the camera, the computer, the software and ultimately, the astronomer. By examining each of these elements in turn, we can quickly rule out some and more painstakingly rule out others, considerably narrowing the avenues of experimentation we need to pursue. We begin with a study of the atmosphere.

2.1 Atmosphere and Gravity Waves

For the purpose of analyzing its thermodynamic properties we can treat the atmosphere as a perfect gas (Scorer 1997). The low density of the atmospheric gas means that the constituent molecules are widely spaced with respect to their size and inter-molecular forces (such as the van der Waals force) are negligible. Gravity binds the atmosphere to the Earth and results in the atmospheric gas settling with the highest density at the

surface. Atmospheric pressure decreases exponentially with altitude as $p = p_0 exp(-z/H)$, where H is the atmospheric scale height (by definition the vertical distance over which the pressure decreases by a factor of e) and is typically of order eight kilometers. The density of dry air at a given altitude is $\rho = P/RT$, where R is the gas constant ($R = 8.314$ J K^{-1}mol^{-1}) and temperature follows a generally linear lapse rate (with different slopes in the troposphere, stratosphere, *etc.*). The standard tropospheric adiabatic lapse rate is defined as $T(z) = T_0 - 0.0065z$ when the altitude, *z*, is given in meters. This results in roughly 80% of the mass of the atmosphere being concentrated in the troposphere (roughly the lowest 15 kilometers of the atmosphere).
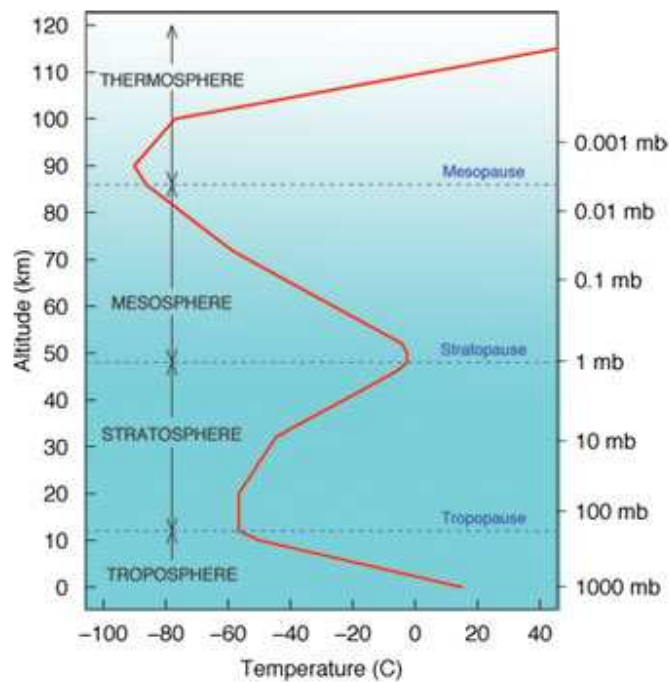
Figure 7. Illustration of atmospheric structure with altitude including temperature (red line) and pressure (right axis). From http://www.learner.org/courses/envsci/unit/text.php?unit=2&secNum=2.

The lowest one to two kilometers of the troposphere are directly influenced by the nature of the underlying terrain and are considered a distinct region called the boundary layer

(see Figure 8). During the day the boundary layer is a highly turbulent region dominated by solar heating and convection. After sunset, the temperature profile of the boundary layer is dominated by terrestrial radiation and cooling which leads to stable stratification where layers of lower density lie on top of higher density layers inhibiting convection. In highly stable conditions, the surface actually becomes cooler than the overlying air and a temperature inversion develops (Stull 1988), meaning that the temperature gradient is reversed so that temperatures increase with altitude in the first few hundred meters of the boundary layer.
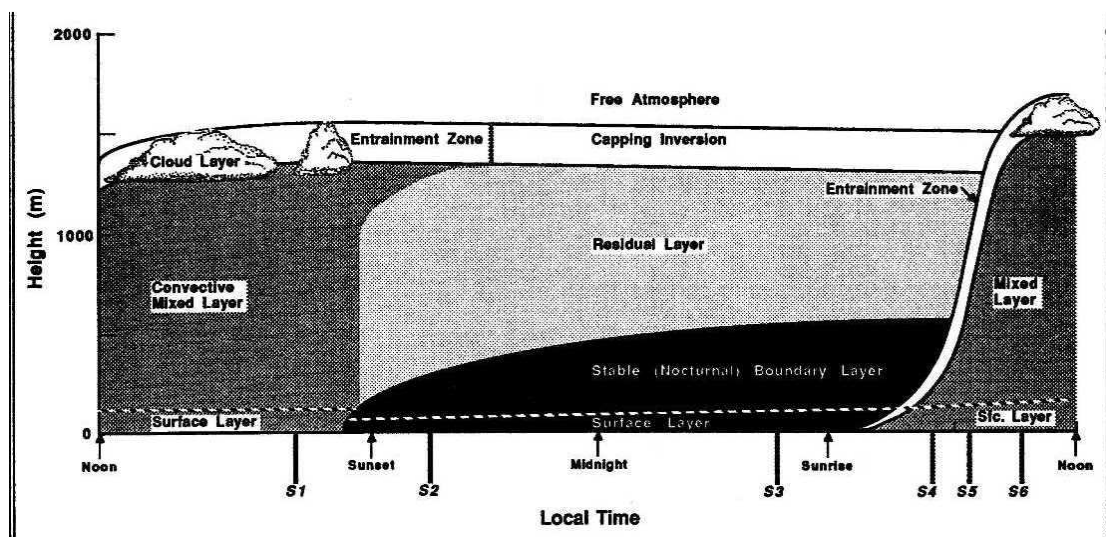
Figure 8. Structure of the atmospheric boundary layer (from Stull 1988).

Winds at different altitudes in the boundary layer can vary significantly in both speed and direction. Near the surface, wind speed is highly dependent on the nature of the terrain – winds immediately above forested areas (*i.e.* high drag regions) may be much slower than those above grasslands or desert. Low-level winds are also influenced by topography. Higher density layers of cold air will flow down-slope resulting in katabatic winds or gravity flows. As height above the surface increases, winds are less influenced by the

21

surface characteristics and approach geostrophic flow, wherein wind speed and direction depend only on a balance of the Coriolis effect and the pressure gradient force. The change in wind speed with altitude is known as wind shear. Another common feature of boundary layer winds is a rotation of the wind direction with altitude due to the Coriolis effect. In the northern hemisphere winds turn clockwise (veer) with altitude while in the southern hemisphere the rotation is counter-clockwise (Haurwitz 1941).

The stability of the atmosphere is quantitatively parameterized by a dimensionless quantity known as the Richardson number. The Richardson number is defined as the ratio of the production of turbulence by buoyancy to that produced by shear (*e.g.* Nappo 2002),

$$R_i = \frac{\frac{g}{\theta}\frac{\partial \theta}{\partial z}}{\left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{dz}\right)^2}. \tag{1}$$

θ is known as the potential temperature and is the temperature a parcel of air would have if lowered adiabatically from a height with pressure p to a height with standard pressure $P = 1000$ mb: $\theta = T(1000/P)^{R/C_p}$, where C_p is the specific heat of air at constant pressure. $\partial u/\partial z$ and $\partial v/\partial z$ are the orthogonal components of the horizontal wind shear. When the Richardson number is large, the work required to displace an airmass from equilibrium is greater than that done by turbulent eddy stresses and the atmosphere is dynamically stable (Haurwitz 1941). At small Richardson numbers, wind shear becomes a dominant force and the atmosphere is dynamically unstable and turbulent.

The key characteristics of the Richardson number are the critical values that signify the boundaries between turbulent and laminar flow. Through a combination of theory and experiment, separate values have been found corresponding to a dynamically stable fluid becoming unstable and the reverse (Stull 1988). Flow transitions from being laminar to becoming turbulent when the Richardson number becomes less than 0.25. Turbulent flow will become laminar if the Richardson number increases to greater than 1.0. The clear hysteresis in the critical numbers comes from the fact that in order for flow to become turbulent, the atmosphere must not only be unstable, but a mechanism must exist to spark the formation of the turbulence. One possible mechanism for the production of turbulence is Kelvin-Helmholtz (K-H) waves, which can only form if the Richardson number is below 0.25. However, if the atmosphere is already turbulent, it can remain so until the Richardson number is greater than 1.0.

The stable nocturnal boundary layer is highly conducive to the formation of atmospheric gravity waves (AGWs). An AGW is formed when a parcel of air in a stably stratified atmosphere is raised or lowered from its equilibrium position. Due to the atmospheric density gradient, an airmass that is lifted from its equilibrium position will have a higher density than the surrounding air and will experience a downward gravitational force. The airmass will then sink, overshooting equilibrium. As it falls into a region of higher density the buoyancy force will result in an upward acceleration and the whole process will repeat. Assuming the airmass has some initial horizontal velocity, an atmospheric wave will be formed. In unstable situations, an airmass displaced above its equilibrium position may encounter a region of even higher density and temperature, thus will

experience an upward buoyancy force and will continue accelerating upwards. In this situation, gravity waves will not form.

From Archimedes's Principle (Archimedes circa 250 BCE) we know that the buoyancy force exerted on a submerged mass is equal in magnitude to the weight of the liquid it displaces. For a volume V_a of air with density ρ_a displaced some distance Δz into a region with density ρ and volume V, the net force is $F = \rho V g - m_a g$, where the mass of the air, $m_a = \rho_a V_a$ and $g = 9.8\ m/s^2$ is the acceleration due to gravity. Because the displaced volume equals the volume of the airmass, the force on the airmass can be rewritten as $F = (\rho - \rho_a)Vg$, which is clearly directed upward when $\rho > \rho_a$ and vice versa. We find the resulting acceleration of a displaced airmass by recognizing that $\rho - \rho_a = (d\rho/dz)\Delta z$ and substituting $\rho_a V a$ for F,

$$a = \frac{g}{\rho}\frac{d\rho}{dz}\Delta z. \tag{2}$$

This equation is analogous to a simple harmonic oscillator with frequency,

$$N = \sqrt{-\frac{g}{\rho}\frac{d\rho}{dz}}. \tag{3}$$

This frequency is known as the Brunt-Väisälä or buoyancy frequency (Wells 1997) and represents the maximum frequency for a gravity wave. From Equation (3) it is apparent that the maximum frequency of a wave is directly related to the density gradient in the part of the atmosphere where the wave is generated. A large density gradient will allow minimal travel of the airmass before it is forced back towards equilibrium, resulting in a higher frequency. In a small density gradient the mass may travel a significant distance before the buoyancy force overcomes gravity, thus creating a low frequency wave.

In the lower troposphere, gravity waves may be caused by forcing of an airmass over elevated topography, wind shear or convection. At higher altitudes wind shear due weather fronts, convective cells, or the jet stream can induce wave activity.

Atmospheric gravity waves in all layers of the atmosphere are a significant player in the study of atmospheric dynamics. Of key interest is the influence of AGWs on energy transport throughout the atmosphere (*e.g.* Fritts & Alexander 2003) and the relationship between AGWs and turbulence (*e.g.* Lu & Koch 2008; Einaudi & Finnigan 1993). A number of recent studies using instrumentation ranging from meteorological towers to Doppler lidars have greatly increased our understanding of the nature of AGWs (*e.g.* Poulos *et al.* 2002; Rees *et al.* 2000; Cuxart *et al.* 2000).
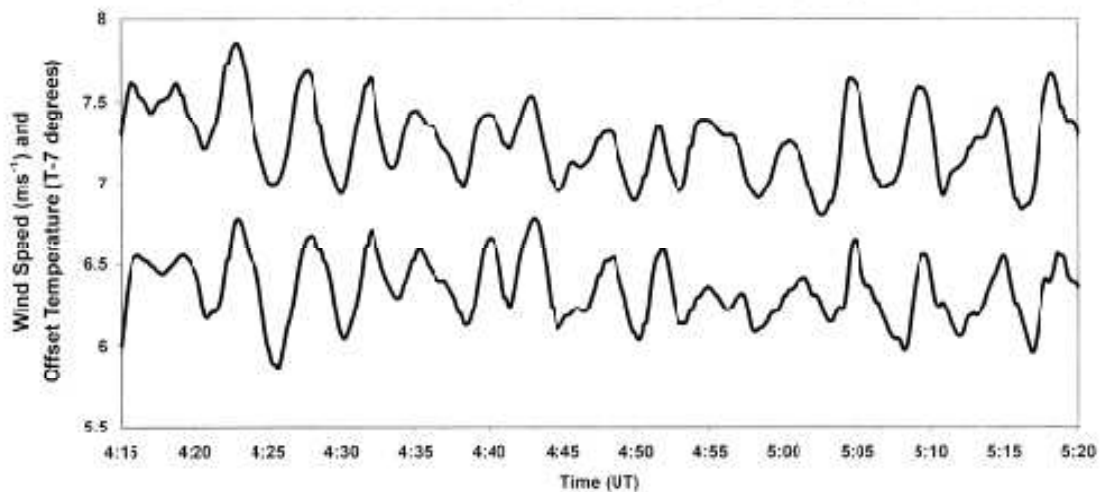
Figure 9. Temperature and vertical wind speed fluctuations due to a gravity wave observed by aircraft at about 470 meter altitude (from Fritts, *et al*. 2003).

During the 1999 Cooperative Atmosphere-Surface Exchange Study (CASES-99), waves were observed in the boundary-layer having wavelengths of 1-10 kilometers and

corresponding temperature fluctuations at the wave altitude of up to one Kelvin (Fritts *et al.* 2003, see Figure 9). Waves in the middle and upper atmosphere will typically have wavelengths of tens of kilometers to global scale phenomena and according to Scorer (1997) we can expect waves to generally have amplitudes smaller than 3 km.

In the boundary layer, or the lowest level of the troposphere where ground interactions are highly relevant, wave characteristics are primarily influenced by the spatial scales of the topography where they are formed (Fritts *et al.* 2003). These waves are commonly observed to have amplitudes of tens of meters to hundreds of meters and wavelengths as long as a few kilometers. In regions where the topography is relatively severe (sharp mountain ridges, tall peaks, significant hills, etc.), we expect to encounter waves with large amplitude to wavelength ratios. Further observations of atmospheric and AGW characteristics will be discussed in Chapter 4.

2.2 Atmospheric Optics

It is appropriate to ask if atmospheric optical turbulence, responsible for the twinkling of stars and familiar to practitioners of adaptive optics, is also responsible for anomalous refraction. It has been suggested that AR is simply a low frequency extension of the atmospheric turbulence spectrum. In the following section we show that this is definitely not the case and that AGWs or other large scale structures are far more likely culprits.

2.2.1 Optical Turbulence

The turbulent nocturnal boundary layer has been a significant source of concern and a major sink for astronomical spending almost as long as the field of astronomy has existed. *Seeing*, the effect of atmospheric turbulence on astronomical observing causes

images to rapidly fluctuate in angular size and position (image motion) with amplitudes (of motion) of order arcseconds.

The seeing quality is often quantified using the Fried parameter or coherence length, r_0, which is defined as the diameter of a seeing cell, or a circular region at the telescope entrance pupil over which the rms phase distortion of an incoming plane parallel wavefront is less than one radian (Roddier 1999). Another way to quantify seeing uses the Strehl ratio, R, which is the ratio of the peak intensity of an observed star to that which would be obtained in a theoretically perfect point source diffraction limited image (for a given telescope). A one radian distortion in the incoming wavefront corresponds to a Strehl ratio of *1/e* and is considered the dividing line between good and poor seeing. For a telescope of diameter smaller than the seeing cell, diffraction limited images are possible, *i.e.* $\theta = 1.22\lambda/d$, where θ is the limiting angular resolution of the telescope, λ is the wavelength of light and d is the telescope diameter. For larger telescopes, each r_0 cell acts as a sub-aperture, such that parts of an incident plane parallel wavefront see numerous optically different "lenses" before reaching the telescope and are randomly refracted into the pupil. The resolution of a telescope with $d > r_0$ is $\theta = 0.98\lambda/r_0$ and is independent of the diameter of the telescope. At most observatory sites, r_0 varies from 0.05 m to 0.2 m, so the atmosphere is the limiting factor in the resolution of nearly all professional telescopes.

The physical structure of optical turbulence and the resulting turbulence spectrum was elucidated through the work of Kolmogorov (1941). In the Kolmogorov model, energy is

injected into large scale turbulent structures, which decay into smaller structures and so forth, with the kinetic energy of the large structure being subdivided into the smaller structures without loss. The size of the largest turbulent structures is called the outer scale, L_0, and is thought to be of order the height above the ground in the surface layer, with typical values of a few tens of meters to hundreds of meters in the free atmosphere. The inner scale, l_0, defines the smallest scales of turbulence and is typically of order a few millimeters (Roggemann & Welsh 1996). Between these bounds, called the inertial subrange, we can quantify the scales and spectrum of refractive index fluctuations by means of the refractive index structure function,

$$D_N(r) = \langle [N(x+r) - N(x)]^2 \rangle = C_N^2 r^{2/3} \tag{4}$$

which represents the mean-squared fluctuation in the index of refraction, N, over a distance, r, $\langle ... \rangle$ denoting an ensemble average. The refractive index structure coefficient, C_N^2, is a measure of the turbulence strength and varies as a function of altitude. In the optical regime index of refraction fluctuations are dominated by inhomogeneities in temperature, such that the temperature structure function follows the same form as Equation (4) with all $N's$ replaced by T. The one dimensional spectral function of the turbulence is the Fourier transform of the structure function and Tatarski (1961) has shown that for a one-dimensional structure function of the form,

$$D_f(r) = c^2 r^p, \tag{5}$$

the corresponding one-dimensional spectral density is,

$$\varphi(\kappa) = \frac{\Gamma(p+1)}{2\pi} \sin\left(\frac{\pi p}{2}\right) c^2 \kappa^{-(p+1)}. \tag{6}$$

Using the fact that the spectral wave number of the turbulence is related to the turbulent cell spatial scale by $\kappa = 2\pi/r$ we find the spectral density function of the turbulence in one dimension is

$$\varphi_T(\kappa) \propto C_T^2 \kappa^{-5/3}. \tag{7}$$

For the purposes of astronomical observing we are concerned with the propagation of light through a three dimensional turbulent medium. According to Tatarski (1961) the three dimensional power spectrum of turbulence is

$$\Phi_T(\kappa) = -\frac{1}{2\pi\kappa}\frac{d\varphi}{d\kappa} \propto C_T^2 \kappa^{-11/3}. \tag{8}$$

We now need to determine how the turbulence spectrum of the atmosphere translates into a noise spectrum in our images. A primary effect of this turbulence is to tilt the plane parallel wavefront being observed, causing variations in the angle-of-arrival and thereby image motion (Hardy 1998). The fluctuations in angle-of-arrival have a Gaussian distribution. The mean-squared image motion over a circular aperture of diameter D is

$$\sigma_m^2 = 0.489 D^{-1/3} \sec\zeta \int_0^\infty C_n^2(h)dh, \tag{9}$$

where ζ is the zenith angle and h is the height above the Earth's surface (Hardy 1998, Greenwood & Fried 1976). Two very significant factors to note are the dependence on telescope diameter – the size of the image motion fluctuations decreases gradually with increasing diameter, and the dependence on the zenith angle. Based on this we would expect the mean-square image motion observed with a one meter telescope to be roughly 20% smaller than that observed by a 0.5 meter telescope at the same zenith angle and under the same conditions. Likewise the mean-square image motion for a given

telescope should be approximately 40% larger at a zenith angle of 45 degrees than at zenith under the same conditions.

The two dimensional spectral density of wavefront tilts (see *e.g.* Hardy 1998, Tyler 1994, Greenwood & Fried 1976 for details of the derivation) is

$$F_\alpha(f) = 0.804 \sec(\zeta)\, v_{-1/3} f^{-2/3} \qquad (10)$$

in units of square radians per Hertz (Hz) for low frequencies ($f \lesssim 1$). At frequencies between about one and ten Hertz the turbulence spectrum falls off as the inverse square of the frequency. Spatial averaging over the telescope aperture at very high frequencies ($f > 10$ Hz) causes the power law to roll over to an $f^{-11/3}$ dependence (Hardy 1998). The turbulence-weighted wind velocity profile is defined at the zenith as $v_m = \int_0^\infty C_N^2(z) v^m(z) dz$, where the velocity is modeled as having a Gaussian profile with height (with a peak at the tropopause). For C_N^2 we used a Hufnagel-Valley profile which gives a value for r_0 of 0.05 m ("HV$_{5/7}$" see Hardy 1998 for a definition and Tyler 1994 for a plot). The power spectral densities of wavefront tilt are plotted as a function of frequency for low frequencies for this particular case of atmospheric conditions in Figure 10.

We can take this analysis one step further by calculating the scales of image motion as a function of frequency for a typical example of atmospheric conditions (Figure 11). The amplitude of the image motion is determined as the square root of average amplitude in a bin (bins are log scale with the exponent increasing linearly by 0.1, so an example of bin is 10^{-2}-$10^{-2.1}$) multiplied by the bin size. Significantly, Figure 11 clearly indicates that the

scales of image motion decrease dramatically from observable motion (~100 mas) at frequencies around one Hertz to motions that are considerably smaller than the observed AR on timescales of tens of minutes.
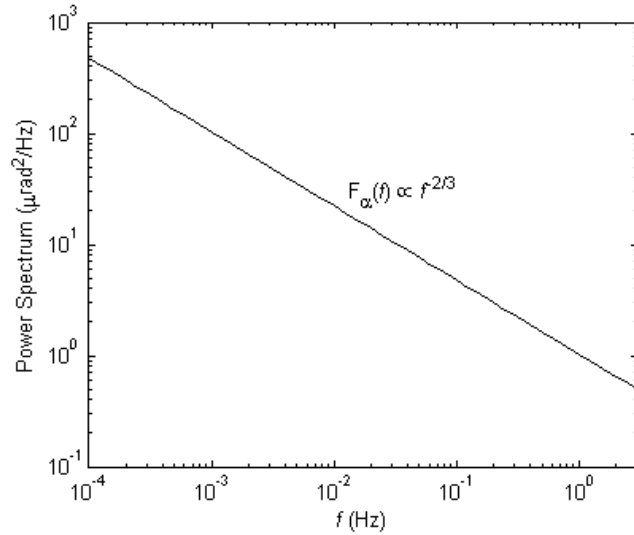
Figure 10. Power spectral density of image motion caused by turbulence-induced wavefront tilt using the $HV_{5/7}$ C_N^2 profile and the velocity profile defined in Hardy (1998).
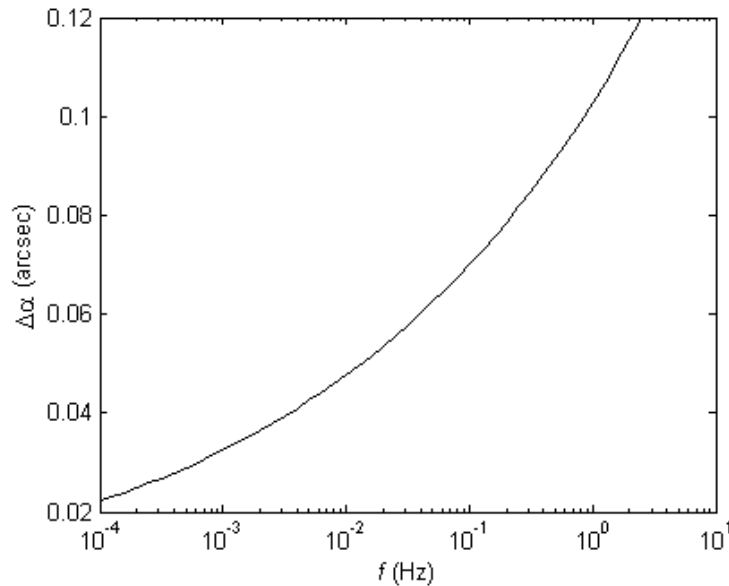
Figure 11. Dependence of image motion amplitude on frequency for the same conditions as **Figure 10**. The dependence will change as frequencies increase above one Hertz due to the steeper power law dependencies of image motion at these higher frequencies.

Because most astronomical images (and all images used in this research) are exposed for timescales of tens of seconds to minutes, stochastic image motions with frequencies of order one Hertz or more will average out over the course of an exposure. This will result in a broadened stellar PSF, but no net displacement of the peak of the PSF. Regardless of the amplitude of the image motion, the only motions that will influence AR are those that produce a net displacement of the stellar image centroid after being integrated for tens of seconds. The image motions that are slow enough to produce a net displacement ($f \leq 10^{-2}$) are nearly an order of magnitude smaller than the observed AR.

Of greater importance to this research than the image motion is the angular scale over which the image motion is uniform – *i.e.* the isokinetic patch. The ubiquitous cell structure of the turbulent atmosphere combined with a finite coherence length ($r_0 \sim 0.05$ m) suggests that the greater the angular separation between points on an incident wavefront, the less coherent their rms distortions on arrival at the focal plane. The scale of coherence of wave phase is called the isoplanatic angle with a phase distortion of one radian considered the boundary between correlated and uncorrelated wavefronts. As a back of the envelope estimate, if we assume that the majority of optical turbulence occurs at or below 10 kilometers and an isoplanatic patch at 10 kilometers is 0.05 meters in diameter (r_0), then the angular scale of that same patch would be roughly one arcsecond. Given this estimate, two stars separated by 10 arcseconds, for example, will have completely unrelated phase. A slightly different approach must be taken to determine the scale of the isokinetic patch, as will be discussed later.

A more rigorous calculation of the anisoplanicity error requires analysis of the phase fluctuations of the incoming plane-parallel wavefront due to layers of turbulence. The phase of the wave observed at a distance L from the turbulent layer is related to the refractive index by $\varphi = k \int_0^L n(z) dz$ (see *e.g.* Roddier 1999, Hardy 1998, Tatarski 1961). From this it can be found that the two-dimensional phase structure function is related to the refractive index structure function by

$$D_\varphi(r) = \langle |\varphi(x_1, y_1, L) - \varphi(x_2, y_2, L)|^2 \rangle$$

$$= 2k^2 L \int_0^\infty \left[D_N\left(\sqrt{z^2 + r^2}\right) - D_N(z) \right] dz,$$

(11)

where $r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$, for $r \ll L$. Substituting Equation (4) into (11) and integrating gives the phase structure function for a single layer of turbulence at altitude L, observed at zenith as $D_\varphi(r) = 2.91 k^2 L C_N^2 r^{5/3}$. For all layers of the atmosphere the structure function is

$$D_\varphi(r) = 2.91 k^2 \int C_N^2(z) dz \, r^{5/3}.$$

(12)

According to Fried (1965), Equation (12) can be rewritten as

$$D_\varphi(r) = 6.88 \left(r/r_0 \right)^{5/3},$$

(13)

where,

$$r_0 = \left[0.423 k^2 \int C_N^2(z) dz \right]^{-3/5},$$

(14)

the value 0.423 having been determined empirically. For observations at a zenith angle ζ, r_0 is multiplied by $(\sec \zeta)^{-3/5}$. If two astronomical objects are observed simultaneously with a separation θ, then at some distance z from the telescope, the observed wavefronts are separated by a length $r(z) = \theta z$. If the source of the turbulence is a layer at an

33

altitude, h, then $z = h \sec \theta$. For multiple layers of turbulence h must be replaced with weighted average $\bar{h}^m = \int_0^\infty dh C_N^2(h) h^m / \int_0^\infty dh C_N^2(h)$ of the layer altitudes (Roddier 1999). Applying r and z to Equation (13) gives the mean-square anisoplanicity error:

$$\sigma_a^2(\theta) = 6.88 \left(\frac{\theta \bar{h} \sec \zeta}{r_0} \right)^{5/3}. \tag{15}$$

Using the definition that an isoplanatic patch is an angular region over which the mean-square phase of the wave varies by less than one radian, we find the angular size of the isoplanatic patch to be

$$\theta_0 = (6.88)^{-3/5} \left(\frac{r_0}{\bar{h} \sec \zeta} \right) = 0.314 \left(\frac{r_0}{\bar{h} \sec \zeta} \right). \tag{16}$$

Using the same conditions as for Figure 10, we find an isoplanatic patch at the zenith of slightly more than 3 arcseconds.

The isokinetic error (or tilt anisoplanatism) is the differential motion of an image centroid due to variation over small angular scales in the tilt terms of the phase distortion of a plane parallel wavefront. The size of the isokinetic patch depends on the telescope aperture and the average height of the turbulence source. According to Kaiser *et al.* (2000) a typical value for the angular scale of the isokinetic patch is one arcminute.

The key significance of this exercise is that for typical astronomical images covering angular scales of tens of arcminutes to degrees, the fraction of the image over which the motion of stars due to seeing is correlated will be very, very small. A single field of view may contain thousands of isokinetic patches with statistically independent motions. In all observations of anomalous refraction to date, the anomalous motions of objects have

34

always been highly correlated across angles as large as a degree or more. This factor is sufficient evidence to firmly rule out optical turbulence (*i.e.* seeing) as the source of AR.

2.2.2 Atmospheric Gravity Waves

As a first approximation an atmospheric wave can be treated as a tilting of horizontal air strata because the wavelength of the AGW is assumed to be very large with respect to the diameter of a telescope and no curvature across a focal plane has been seen in anomalous refraction data. (The scale of curvature in atmospheric tilt has been examined with Sloan telescope data as part of this research and found to be sufficiently small for the tilt approximation to be valid, as will be discussed in Chapter 3.) We parameterize the wave by assuming it is an infinitely thin plane interface between two atmospheric layers with different indices of refraction which are tilted by some angle α (see Figure 12). The tangent of this angle is an approximate function of the amplitude of the wave and the wavelength. Parallel light rays entering the atmosphere from the zenith will encounter the tilted interface at an angle α with respect to the normal and will be refracted toward the normal because they are traveling from a medium with lower index of refraction to one of higher index.
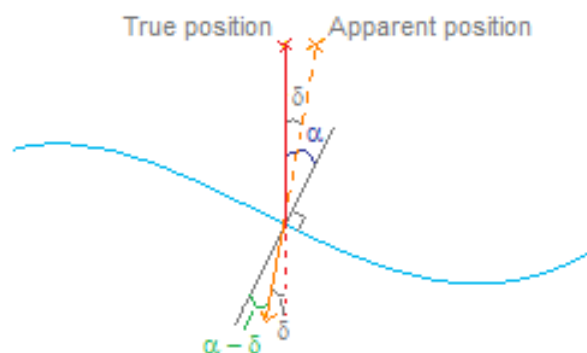
Figure 12. Refraction effect of an atmospheric gravity wave approximated as a tilted plane interface between strata with different indices of refraction.

Using geometric optics, it is a straightforward matter to determine the refraction of a zenith ray due to an AGW in this model. From Snell's law we find the relationship between the respective indices of refraction of two media and the angle of refraction of a light ray passing from one to the other is, $n_1 \sin \theta_1 = n_2 \sin \theta_2$. In the case of the tilted plane interface approximation, $\theta_1 = \alpha$ and $\theta_2 = \alpha - \delta$. Using the fact that $\sin(\alpha - \delta) = \sin(\alpha)\cos(\delta) - \cos(\alpha)\sin(\delta)$ and dividing through by $\cos(\alpha)$ we get

$$\tan \alpha = \frac{n_2 \sin \delta}{n_2 \cos \delta - n_1}.$$ (17)

Because the zenith refraction angle, δ, is quite small (based on the observations of anomalous refraction mentioned in the introduction, we expect δ to be of order a few tenths of an arcsecond), the above equation can be simplified and rearranged to give the angle of refraction from the zenith (anomalous refraction) as a function of the tilt angle and change in index of refraction

$$\delta = \frac{n_1 - n_2}{n_2} \tan \alpha.$$ (18)

The slope of the wave at any point (tan α) is determined by taking the derivative of the wave equation with respect to position. Because the actual angle we are concerned with is the angle an incoming ray makes with the normal to the wave at any point (90 degrees counter-clockwise from the slope angle, see Figure 12), we actually want the negative of the derivative (such that a wave with negative slope and normal pointed to the right of zenith has a positive tilt angle).

36

Lifting an airmass above its equilibrium position will have the effect of changing the temperature profile of the atmosphere. Instead of the temperature being constant along a horizontal layer of the atmosphere, the temperature will be raised at the horizontal position of the peak of a wave and lowered at the position of the trough. In other words the isothermal surface is tilted by the passage of the wave.

The refractivity of moist air is calculated based on work by Ciddor (1996, see reference for full details of calculation and Appendix A, code refractivity.m, for the full calculations used in this research). For dry standard air (zero percent humidity, temperature of 15° C, pressure of 101,325 Pa and 450 parts per million (ppm) CO_2) the refractivity is

$$(n_{as} - 1) \times 10^8 = \frac{k_1}{(k_0 - \sigma^2)} + \frac{k_3}{(k_2 - \sigma^2)}, \tag{19}$$

where σ is the wave number in inverse microns and the constants are described in Appendix A. Non-standard values (x_c) for the carbon dioxide modify the refractivity as,

$$(n_{axs} - 1) = (n_{as} - 1) \times [1 + 0.534 \times 10^{-6}(x_c - 450)]. \tag{20}$$

Water vapor at the standard conditions of 20° C and 1333 Pa has a refractivity of

$$(n_{ws} - 1) \times 10^8 = 1.022(w_0 + w_1\sigma^2 + w_2\sigma^4 + w_3\sigma^6), \tag{21}$$

where the constants w_i are defined in Appendix A. From Equations (20) and (21) we find the total refractivity of moist air under experimental conditions to be

$$(n - 1) = \frac{\rho_a}{\rho_{axs}}(n_{axs} - 1) + \frac{\rho_w}{\rho_{ws}}(n_{ws} - 1), \tag{22}$$

where the first term on the right side corresponds to the dry component of moist air with standard density, ρ_{axs} and observed density, ρ_a, while the second term on the right is the water vapor component with standard density, ρ_{ws} and observed density, ρ_w.

Because any pressure changes caused by a wave or turbulent atmospheric motions at the altitude of the wave are negligible compared to the corresponding temperature changes (Scorer 1997), variability of the refractivity of the atmosphere at a given altitude and wavelength due to an AGW can be treated as a function of the temperature change alone. This becomes clear when we consider that density, a key component of the index of refraction, is proportional to pressure divided by temperature, where pressure is in Pascals and temperature is in Kelvin. Standard atmosphere sea level temperature is 288 K while standard pressure is 1013.25-mb. A reasonable temperature fluctuation due to a wave would be one Kelvin (0.35%), so for the pressure change to have an effect comparable to that of the temperature, the fractional change would have to be greater than or equal to 0.35%, or 3.5-mb. This is the sort of pressure change you might see over the course of a day, but is orders of magnitudes greater than what a wave might induce.

The change in refractivity of the air and refraction of light from a zenith star is plotted as a function of atmospheric gravity wave parameters in Figure 13. The top two plots show us that the effect of a temperature change on the refractivity, μ, at a given altitude is several orders of magnitude greater than the effect of a change in water vapor mixing ratio (*r,* typical values of order 10 g/kg). Increasing the altitude decreases the change in the refractivity for a given temperature change, with a dramatic fall off in $\Delta\mu$ above the

troposphere (*i.e.* above ~11-km). The refraction of the light from a zenith star is plotted as a function of atmospheric tilt and change in the refractivity in the bottom plot. Changes of refractivity of order 10^{-7} or smaller seem unlikely to cause refractions comparable to those observed unless the perturbed atmospheric layers are tilted by considerably more than 45 degrees from the horizontal. It is clear from this figure that the primary source of undesirable refractions will be waves in the lowest few kilometers of the atmosphere, with the same temperature differences having increasingly smaller effects at higher altitudes.
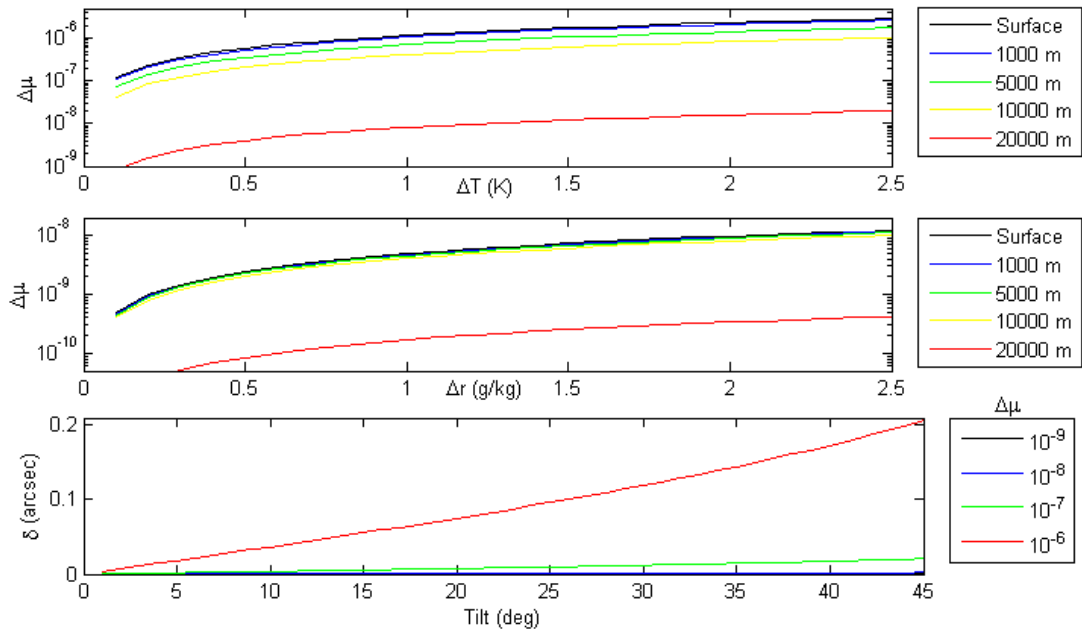
In addition to the reduced changes with altitude in index of refraction due to AGWs, recall that AGWs in middle and upper atmosphere generally have long wavelengths (tens

to thousands of kilometers) and are unlikely to have the correspondingly large amplitudes to produce the atmospheric tilts necessary for an observable refraction.

Although simplistic, this model indicates that a single tilted interface between atmospheric strata of differing refractivities at low altitude can cause a refraction of the same magnitude as the anomalous refraction observations discussed in the first chapter. An example of a wave at a one kilometer altitude which would cause a 0".1 refraction is one that can be parameterized by a tilted interface angle of 25 degrees and a temperature difference of 0.75 K. An idealized wave with its associated atmospheric tilt and resulting refraction are modeled in Figure 14.
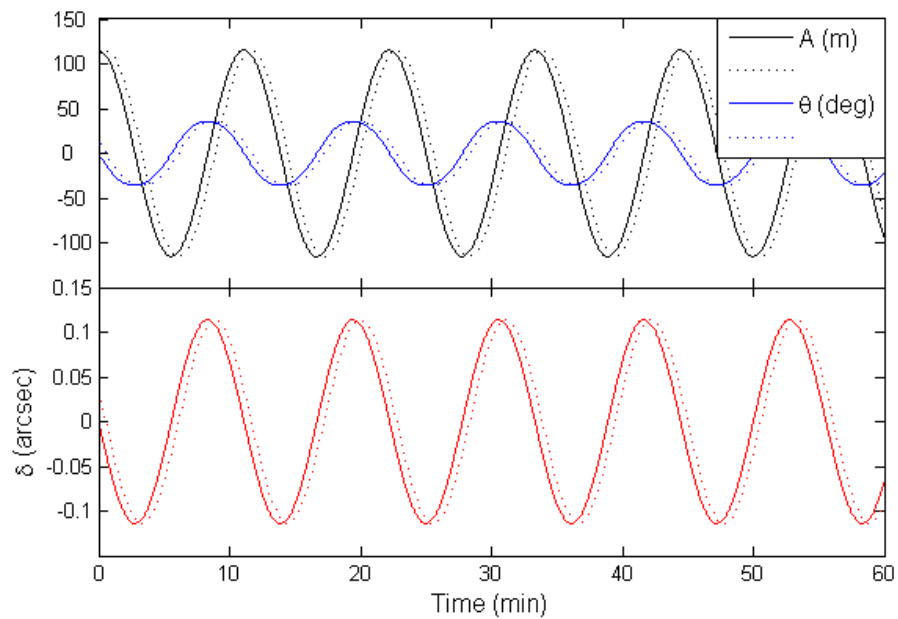
Figure 14. Example of an AGW 500 meters above an observatory at 1000 meter elevation. The wave has a 1000 meter wavelength, 115 m amplitude and a 0.75 K temperature amplitude, and is traveling at 1.5 m/s. The bottom plot shows the resulting refraction and the dotted lines indicate the wave and refraction as viewed from a second location 50 meters from the first.

40

With a number of new very wide field-of-view (FOV) stare-mode telescopes in production whose primary science drivers include astrometry (*e.g.* LSST, Tyson 2002, and PanSTARRS, Kaiser, *et al.* 2000), it is imperative to understand how anomalous refraction may affect their observations. In stare-mode relative astrometry, a constant refraction across the field of view (DC offset, effectively) is easily modeled when comparing images to reference stars. What is of significantly more concern is an unknown systematic refraction which is variable across the field of view and relatively constant over the duration of the image. Anomalous refraction as observed in TDI observations varies on timescales of minutes to tens of minutes, but as the source is hypothesized to be atmospheric gravity waves, a wide field "snapshot" of anomalous refraction would show a smoothly varying refraction across the field of view which would appear stationary on short timescales. To understand how this effect might appear to wide FOV telescopes, the refraction caused by a wave viewed at high zenith angles across a wide field of view was modeled.

This model simulates an isolated single frequency AGW at an altitude of 500 m above a 1000 m elevation observatory under standard temperature and pressure (same model as depicted in Figure 14). The wave is observed by a telescope with a 4° FOV at a zenith angle of 45°. Because the wavelength of the wave is large with respect to the FOV of the telescope, at each point the wave is modeled as a plane surface which is tilted with respect to the horizontal. The coordinate system used is illustrated in Figure 15. Zenith is at 0° and angles increase towards the east. In the model, wave tilts are measured as the

41

angle the normal of a point on the wave makes with the zenith. A negative refraction is that which makes an easterly star appear higher in the sky. As modeled, the telescope is looking towards the east.
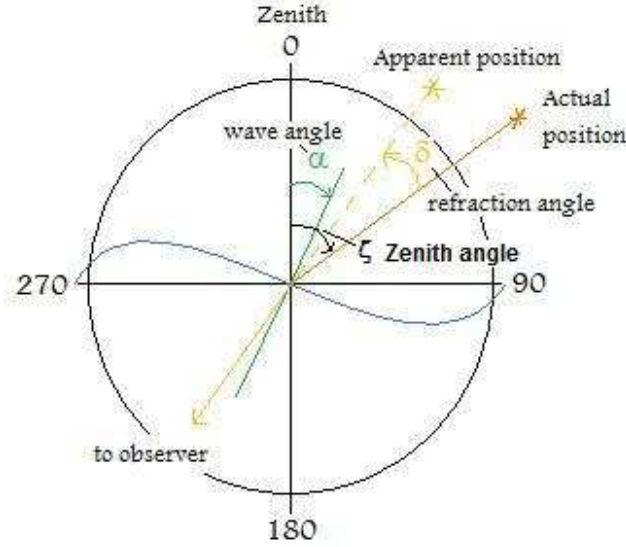
Figure 15. Coordinate system used in wide-field model.

The refraction due to observing at non-zero zenith angles through a plane parallel atmosphere is

$$R_{n_0} = R_0 \frac{PT_0}{P_0 T} \tan z_t \tag{23}$$

where z_t is the true zenith distance, P and T are the atmospheric pressure and temperature with the 0 subscript indicating standard conditions and the constant of refraction is $R_0 = (n_0^2 - 1)/2n_0^2 = 60''.35$, with n_0 indicating the standard atmosphere index of refraction.

The results of the model are shown in Figure 16. These results indicate that anomalous refraction may diverge from the standard zenith angle refraction by more than an arcsecond, with the difference in the anomalous refraction (with standard zenith angle

42

refraction removed) between opposite ends of the field of view varying from zero to several tenths of an arcsecond on timescales of the wave period. Adding further layers of waves above and below those modeled here, as well as curvature of atmospheric strata across the focal plane will likely result in much less predictable refraction.
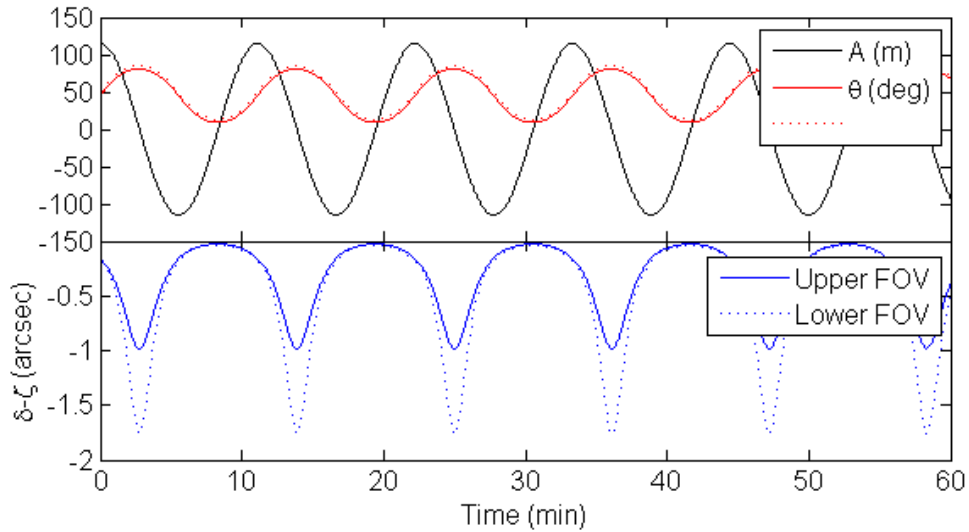
Figure 16. Same wave as **Figure 14** observed by a telescope with a four degree field of view from a 45° zenith angle. Solid and dotted lines indicate wave angle (upper figure) and anomalous refraction, δ, minus zenith angle refraction, ζ, (lower figure) as viewed from upper and lower ends of the FOV respectively.

An example of where this may become a significant concern is using a very wide field telescope like the LSST to create a mosaic of the sky on short timescales. The observing strategy of the LSST is to observe a large fraction of the sky several times per month using very short (10 second) exposures (Tyson 2002). In a single image LSST would see significant variation in anomalous refraction across the 3° FOV as indicated by Figure 16. Adjacent images taken at different times would have unrelated errors due to AR associated with them. To create a mosaic, overlapping stars at the edges of adjoining images are matched to determine relative image placement. AR will effectively distort

43

each frame in two dimensions making aligning images to high precision difficult if not impossible.

Based on this study of the theory of atmospheric gravity waves and their hypothesized relationship to anomalous refraction it seems entirely plausible that these waves are a source (if not *the* source) of anomalous refraction. The waves which will be of most concern to us are those which occur in the lowest levels of the atmosphere, typically in the lowest few kilometers above the surface and which have the largest amplitude-to-wavelength ratios and most significant temperature differentials at the wave altitude. However, waves with very small temperature differentials but large amplitude to wavelength ratios will also play a large refractive role. When these waves occur depends on the stability of the atmosphere, and the wave characteristics are largely determined by the wind and topography conditions under which they are formed.

2.3 Other possible sources of AR

In order to fully cover the topic of anomalous refraction it is necessary to consider any and all other possible sources of the effect. In continuing with our examination of sources along the optical path we will first consider alternate atmospheric sources followed by a brief discussion of instrumental and observer issues.

2.3.1 Coherent turbulence structures (ramps)

Atmospheric physicists studying boundary layer turbulence (*e.g.* Antonia *et al.* 1979, Cava *et al.* 2004) have noted the occurrence of coherent structures known as "ramps" occurring in the stably stratified nocturnal boundary layer. These ramps, generally observed with temperature probes and anemometers on a meteorological tower, appear as

an abrupt increase in temperature at a given altitude followed by linear decrease over a period of tens of seconds to a few minutes (see Figure 17). The temporal scales of ramps are generally smaller (typically less than one minute periods) than has been observed in astrometric data; however, we still need to consider whether ramps could contribute to anomalous refraction.
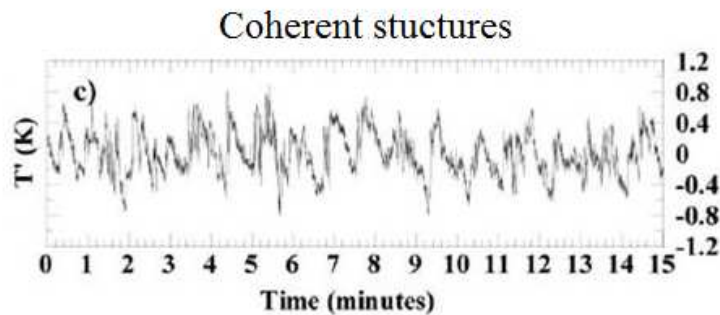
Figure 17. Time series observations of the temperature changes characteristic of ramps over a forest canopy, observed using an instrumented tower (adapted from Cava *et al.* 2004).

To understand the refractive nature of ramps we modeled them as a simple sawtooth wave described by $y = A \times \text{frac}(x/T + \phi)$, where A is the amplitude, T the period, ϕ the phase and $\text{frac}(x) = x - \text{floor}(x)$. The refraction of a zenith ray is determined based on the angle of tilt of the interface between two indices of refraction as for atmospheric gravity waves using Equation (18). A modeled ramp waveform and the resulting refraction are plotted with respect to time in Figure 18. It is clear from this figure that ramps are not a source of anomalous refraction. The constant angle of atmospheric tilt maintained during most of the ramp cycle will result in a small but constant refraction. (The refraction jumps in Figure 18 are an effect of the model used and are not expected in real atmospheric refraction situations.) Because we are primarily concerned with

~minute timescale quasi-periodic behavior in this research, we can disregard ramps as a significant effect.
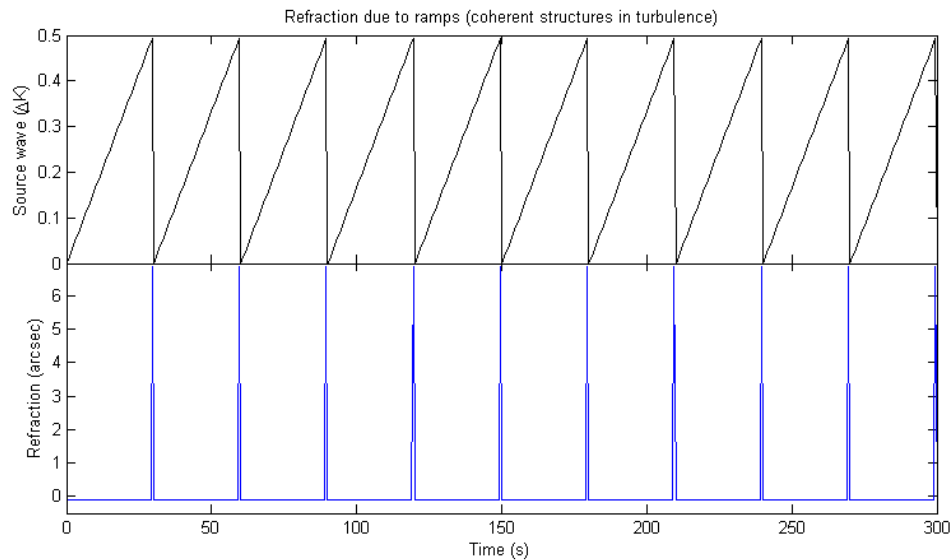
Figure 18. Modeled atmospheric ramps in the boundary layer and the resulting refraction of a zenith ray. Jumps in refraction are due to the model used and are not expected in real atmospheric refraction conditions.

2.3.2 Canopy Waves

Where terrain is characterized by either dense or sparsely populated forest the nocturnal boundary layer within a few tens of meters of the canopy top will frequently be perturbed by a type of atmospheric gravity waves called canopy waves. Canopy waves result from wind shear (the air beneath the canopy top typically being calm while air above the canopy may have large horizontal velocity) and have spatial characteristics that are dependent on the density of the canopy (Brunet & Irvine 2000). Forests with closely packed trees have greater wind shear immediately above the canopy top than their sparsely populated counterparts. Cava *et al.* (2004) observed both canopy waves and ramps using an instrumented tower extending just above a pine forest (Figure 19).

46

Canopy waves were observed to occur during periods of high atmospheric stability and low winds, while ramps dominated during higher wind conditions.
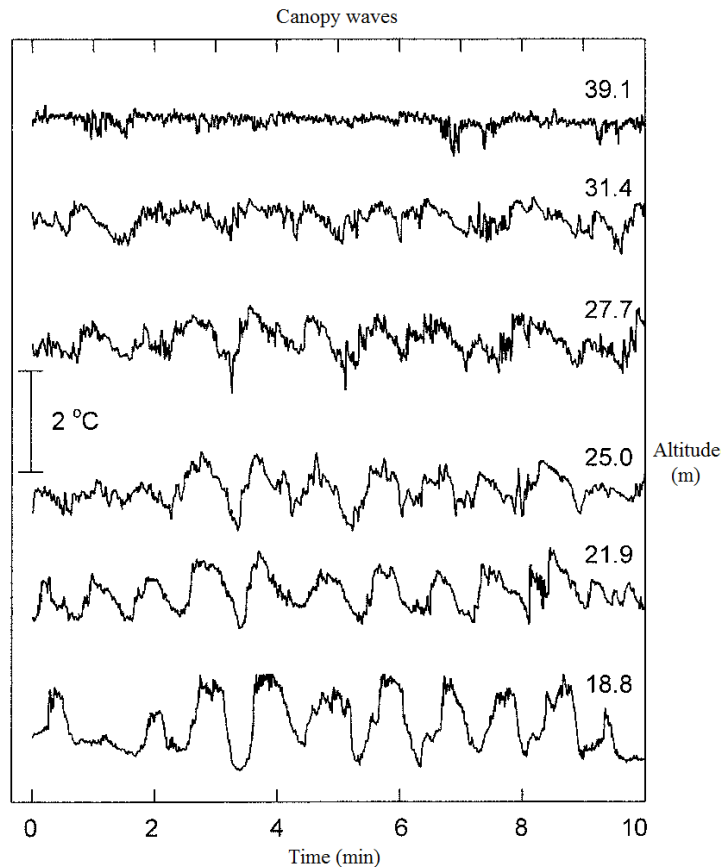
Figure 19. Time series of temperature changes due to waves above a forest canopy, observed with an instrumented tower (adapted from Lee 1997).

Canopy waves are generally described as having short wavelengths (with respect to higher altitude atmospheric gravity waves) of order 100 meters and times scales of one or two minutes. These waves tend to have a single dominant frequency at any given time and may only last a few wave cycles. The short periods and evanescent nature of these waves, coupled with their inherent ties to forested terrain (not a universal observatory characteristic) lead to the conclusion that canopy waves are an unlikely source for anomalous refraction.

47

2.3.3 Kelvin-Helmholtz Billows

When the atmosphere is moderately unstable (*i.e.* $R_i < 0.25$) wind shear can cause the formation of Kelvin-Helmholtz billows. These structures begin as gravity wave-like perturbations of the sheared layers, but due to the unstable nature of the atmosphere, their amplitudes rapidly increase until the waves crest, roll up and form vortices which dissolve into turbulence. This is a transient phenomenon because once the waves begin to roll up, higher density layers become displaced above lower density layers resulting in an instability that causes the waves to break down (Nappo 2002). The time period over which this evolution occurs is usually between 15 and 30 minutes (Blumen *et al.* 2000).

K-H billows have been observed throughout the boundary layer (*e.g.* Figure 20, Chimonas 1999; Newsom & Banta 2003) and are considered one of the predominant sources of mixing in a stratified medium (Scorer 1997). These billows occur in trains of several waves with wavelengths ranging from as small as a few meters to less than a kilometer and the ratio of the billow amplitude to the wavelength is typically less than one (Blumen *et al.* 2000; Chimonas 1999; Scorer 1997). Wave periods are of order minutes or less.
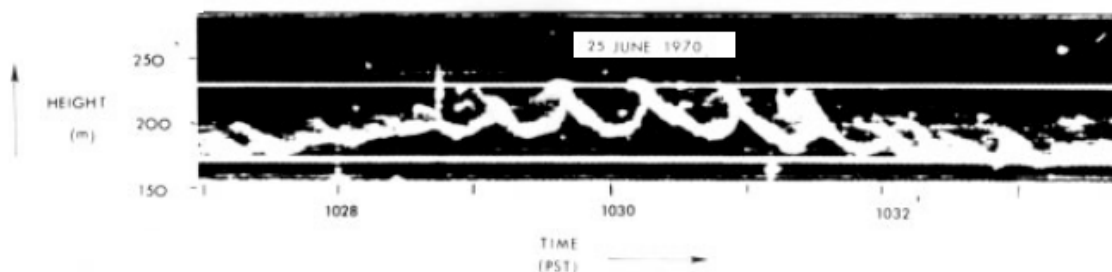
Figure 20. Kelvin-Helmholtz instability imaged with a Frequency-Modulated Continuous-Wave (FMCW) Radar (from Chimonas 1999).

The complicated and evolving non-linear nature of Kelvin-Helmholtz waves made them much more difficult to model than the other studied atmospheric dynamics. For the actual simulation of the temperature field of an atmosphere perturbed by a K-H instability, I used a MATLAB program designed by Danaila *et al.* (2005). The program solves the two-dimensional Navier-Stokes equations for a fluid of density ρ_1 sandwiched between two layers of fluid with different density ρ_2 modeled on a two-dimensional grid. The grid is initialized with a *u* (velocity along iso-density layers) velocity gradient that has a peak at *y = 0* (center line of grid) and decreases towards the bottom and top edges (this simulates a jet of higher density fluid entering a lower density medium). K-H billows form at the top and bottom edges of the jet and are visualized both through vorticity and a passive scalar which can represent any characteristic of the fluid that traces the dynamical structure, but does not influence it (such as temperature or aerosols).

For the purposes of deriving an astronomical refraction, we treated the passive scalar as the temperature field of the fluid and chose a temperature range of two Kelvin across the disturbance in agreement with the temperature characteristics of K-H billows observed by Blumen *et al.* (2000). The dominant wavelength and phase speed were also based on the observations of Blumen *et al.* with respective values of 320 m and 5.5 m/s. In order to approximate a boundary-layer wave above a high altitude observatory the pressure altitude is set to 1500 meters. Because the simulation actually models a jet, only the top half of the grid (which is independent of conditions in the bottom half of the grid) is necessary to simulate a single layer K-H instability due to a velocity gradient. The resulting Kelvin-Helmholtz wave is described by a temperature field as a function of time

49

and altitude. At each point (i, j) on the grid (i increases in the positive time direction and j increases with increasing altitude), the tilt of the atmospheric strata is defined as $\tan \alpha = \frac{T(i+1,j)-T(i,j)}{dT/dz \times \Delta x}$, where dT/dz is the vertical temperature gradient and Δx is the horizontal distance between points $(i+1, j)$ and (i, j). The index of refraction at each point, $n(i, j) = n(P, T(i, j), r)$, where P is the pressure altitude and r is the water vapor mixing ratio, is calculated using Equations (19) through (34). The cumulative refraction at each altitude and time of a ray entering the top of the atmosphere at zenith is then determined using a modified form of Equation (18),

$$\delta(i,j) = \delta\big(i, j_{max}:(j+1)\big) + \frac{n_2 - n_1}{n_2}\left(\frac{\tan \alpha - \delta(i, j+1)}{1 + \delta(i, j+1)\tan \alpha}\right), \tag{24}$$

where $\delta\big(i, j_{max}:(j+1)\big)$ is the cumulative refraction due to all layers above the atmospheric layer in question. The described Kelvin-Helmholtz wave and associated refraction are shown in Figure 21.
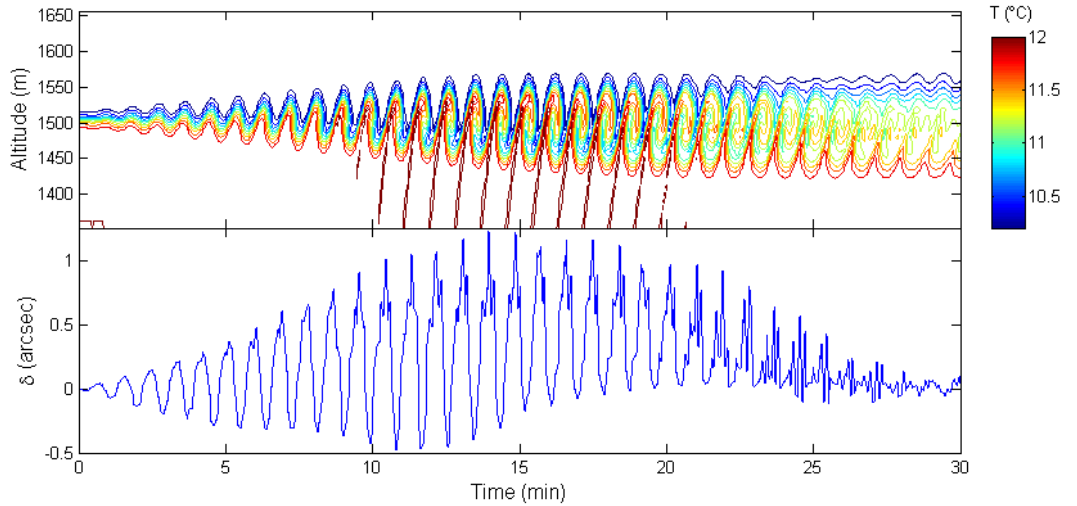
Figure 21. Modeled Kelvin-Helmholtz instability with 320 m wavelength, 2 K temperature difference and phase speed of 5.5 m/s (top), and resulting anomalous refraction (bottom).

Based on the results of this model, Kelvin-Helmholtz billows appear to be another viable source of anomalous refraction. The refraction amplitudes are easily large enough to compare with observations, and the periods fall inside the short end of the minute to tens of minutes period range described in published AR accounts.

2.3.4 Microseismic Activity

Moving to terrestrial sources, one consideration is microseismic motions of the Earth beneath the telescope. AR could conceivably be a result of solid body tilting of the telescope structure (and its environs). Any strictly vertical motions of the telescope will not affect our images, so we can treat tilting of the Earth's surface as a rotation about an axis centered on the telescope. Because there is a direct one-to-one translation between the amount of telescope tilt and the degree of apparent shift of stars in the focal plane, the resulting residuals should have the same characteristics regardless of the nature (structure, size, focal length, *etc.*) of the telescope with which they are observed. If we imagine having observed a generous maximum residual displacement of one arcsecond, this would correspond to a telescope tilt of also one arcsecond, or roughly 5×10^{-6} radians. In terms of strain, we would require a vertical surface deflection of about 5 microns per meter horizontal span for the same arcsecond tilt and these deflections would need to occur with frequencies of 10^{-4} to 10^{-2} Hz to match our observations. Laser strain-meter data taken over the course of three years at several well spaced locations across the continental US (Berger & Levine 1974) indicates the range of seismic amplitudes and frequencies that we could expect to experience at any continental observatory site (see Figure 22). Taking the square root of the average power spectral density in the frequency range of 10^{-4} to 10^{-2} Hz (call it 10^{-19}/Hz) gives us a strain

amplitude of 10^{-4} microns per meter per frequency bin (10^{-10} meters per meter per Hz),
which would correspond to an apparent image motion of 0.01 *milli*-arcseconds. If we
assume that the scale length of Earth strain is greater than the size of an observatory, say
one kilometer, such that the entire observatory responded to ground motions as a solid
body, the total deflection amplitude across the observatory would be of order 0.1
microns. Even in the extreme case of a magnitude 7.5 earthquake, the same scenario
would only result in surface deflections of 10 microns per kilometer per Hz in the
frequency band of concern (granted, the rather more significant motions at higher
frequencies would make observing generally inadvisable anyway). It is clear from this
analysis that Earth motions due to microseismic activity are orders of magnitude too
small to result in any noticeable positional errors in our astrometric data.
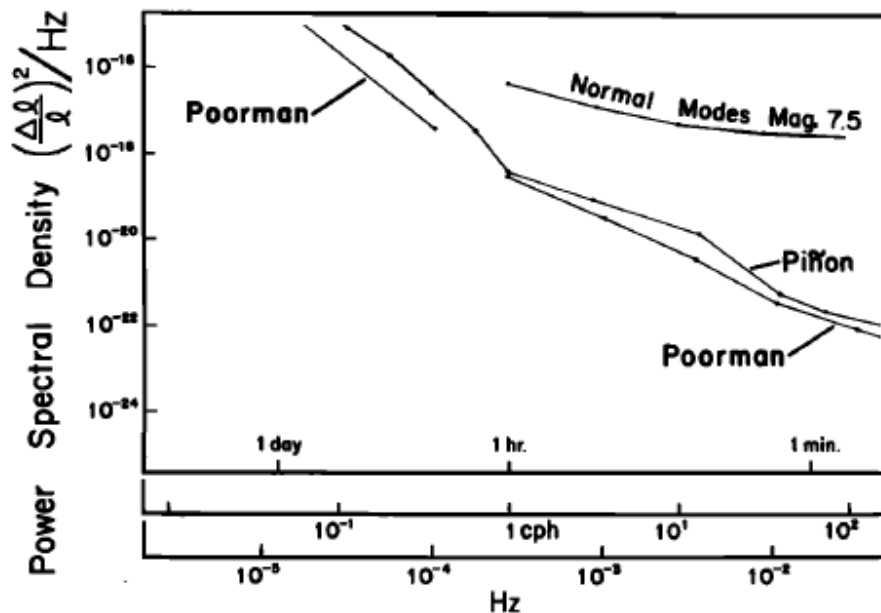
Figure 22. Power spectral density of Earth strain from three years of laser strain meter data. The upper line labeled Normal Modes Mag. 7.5 represents spectral levels during an earthquake (adapted from Berger and Levine 1974).

2.3.5 Instrumentation and Observers

The final elements in the "optical path" that need to be considered are the telescope itself, the camera, data processing and astronomers. The most immediately obvious source of uniform motion of a focal plane is motion of the telescope itself. If telescope motion was the source of AR, we would expect that AR observed using radically different telescopes would display different period and amplitude characteristics. In reality the anomalous refraction described by observers using the 8" FASTT (a meridian transit instrument with only one axis of motion, Stone *et al.* 2005) is characteristically the same as the anomalous refraction observed with the 2.5 meter Sloan telescope (a modified Ritchey-Cretien telescope design, Pier *et al.* 2003). Along the same lines, Schlesinger observed the same sort of AR using the 40" Yerkes refractor (1916) that Hudson recorded using two cameras (1929). The characteristic consistency between all of these observations made not only on widely varying telescopes, but also with highly contrasting methods of record: the naked eye, photographic plates or CCDs, leads to the conclusion that the camera used (or lack thereof) is also not the source of the AR. This also rules out the data processing methods because these have ranged from recording positions by hand to measuring photographic plates to centroiding pixels. And finally the sheer number of observers who have encountered anomalous refraction, many of whom are or were prominent astrometrists, suggests that anomalous refraction is not an observer error.

2.4 Observational Expectations

Anomalous refraction is believed to be caused by the optical influences of the passage of AGWs over the telescope; which is why it is referred to as anomalous *refraction*.

Considering this hypothesized source, we can immediately consider several characteristics that should be present if this is an atmospheric refraction effect.

Multiple observations made by different telescopes simultaneously at the same site, viewing roughly the same region of sky should exhibit similar residuals once telescope motion is subtracted. Likewise, residuals from all telescopes should show similar characteristics regardless of the telescope design or construction. Early experiments using side-by-side visual transit instruments (Schlesinger 1905) and adjacent cameras capturing star trails (Hudson 1929) were able to observe this similarity on very long timescales (annual) and very short timescales (one minute). We seek to confirm these observations using more modern techniques.

The intrinsic properties of optical refraction point toward additional characteristics that should be present in anomalous refraction. It is a basic optical principle that any refraction effect will experience dispersion. Thus we expect residuals from images made with a given telescope using different wavelength filters should consistently show a difference in amplitude of the refraction in accordance with the color dispersion of the atmosphere. Based on Equations (19) through (22), we predict that the peak-to-peak amplitude of a particular anomalous refraction event observed in a 353 nm (u') filter would be approximately 5% larger than the amplitude of the same event observed through a 835 nm (z') filter. In other words, an anomalous refraction that has a 0.5 arcsecond amplitude when observed through a 835 nm filter will appears as a 0.525 arcsecond deflection with a 353 nm filter.

If anomalous refraction is caused by traveling atmospheric waves, we expect to see a phase lag in the periodic residuals of images taken by CCDs at opposite ends of a large field of view or by telescopes separated by some distance viewing the same field on the sky. If the two CCDs view fields on the sky separated by one degree, a given phase of a wave passing overhead at a thousand meters will travel approximately 17.5 meters between passing over the first and last CCD. For a thousand meter wavelength wave, this amounts to a phase lag of a little less than 2% of the wavelength (see Figure 23). If the wave had a ten minute period with respect to the surface (a phase speed of 1.67 m/s), a given phase would take about ten seconds between passing over one CCD and passing over the other.
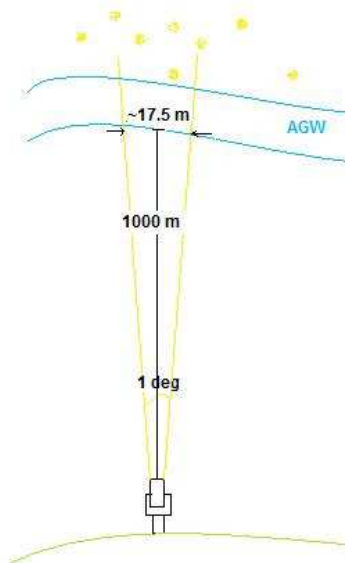
Figure 23. Example of how a phase lag could occur between anomalous refraction observations made by CCDs on opposite ends of a large field-of-view.

The final test of the atmospheric wave hypothesis for anomalous refraction is to look for the direct correlation between observed gravity waves and observed anomalous

55

refraction. If caused by AGWs, anomalous refraction can only be present when conditions allow AGWs to also be present (*i.e.* the atmosphere is stably stratified). Additionally, simultaneous telescope observations and AGW observations should show that anomalous refraction occurs simultaneously with atmospheric waves and that they show corresponding periods, amplitudes, etc.

In the chapters that follow we will investigate each of these qualifications. Based upon the often repeated hypothesis that AR is caused by AGWs, if we can show that anomalous refraction meets all observational tests, then we can be safe in our assumption that this effect is caused by atmospheric gravity waves. The broader investigation beyond confirming the reality and origins of AR will include investigating all the possible sources of AR (not already ruled out) described in this Chapter.

3. Existing Data on Anomalous Refraction

We begin the observational phase of this research with analysis of existing astrometric data from both the original CCD/Transit instrument (CTI) and the Sloan Digital Sky Survey (SDSS). We also examined atmospheric data from the Atmospheric Radiation Measurement (ARM) archives, including data from the Atmospheric Emitted Radiation Interferometer (AERI).

With the CTI data, we were primarily attempting to determine whether AR is actually present in the dataset in support of the hypothesis that AR is universal to all ground-based astrometric observations. The size and characteristics of the SDSS dataset made it ideal for answering a number of questions regarding AR. Specifically, we looked at whether AR is a continuously occurring or occasional phenomenon, and if the latter, what conditions are associated with its occurrence; whether AR is correlated across the SDSS field of view, and if so whether dispersion or phase lags can be seen in the residuals. The analysis of the SDSS data also allows us to corroborate Pier's (2003) assessment of AR in the same data.

The atmospheric analyses were suggested by an atmospheric physicist who suggested that AERI data might be relevant for this research (Nasiri 2009, private communication). This dataset includes several years of lower tropospheric (up to 3 km) measurements of temperature and water vapor profiles. While this data is unrelated in time and location with any of our astrometric data, it provides us with an extended collection of moderately high time and altitude resolution atmospheric observations and provides data on typical

boundary-layer and lower troposphere conditions. We also extended our theoretical atmospheric refraction analysis to this dataset to examine the typical refractions we might expect from real atmospheric conditions.

3.1 CTI

CTI was the original drift-scan telescope, and drift-scan, or Time-Delay and Integrate (TDI) CCD readout mode is the best astronomical observation mode for observing anomalous refraction because it resolves AR in the time domain. Because all astronomical data used during this research were taken in TDI mode, prior to discussing the data and observations we will discuss the important details of TDI operation. We first consider the nature of a CCD and standard CCD operation.

A CCD (Charge Coupled Device) is a two-dimensional array of semiconducting silicon pixels. In a common three-phase CCD each pixel is defined by three electrodes (A, B and C) attached to three separate voltage supplies (with a single voltage supply controlling all of the A electrodes, for example, on a CCD) with channel stops implanted between columns (see Figure 24). Light incident on a semiconductor excites electrons into vacant states in the conduction band leaving an equal number of holes in the valence band. The electrons in the conduction band are free to move throughout the material under the influence of a potential. To produce an image, a potential well is created in each pixel location by maintaining all of the B electrodes at an appropriate (positive) voltage during the exposure while the A and C electrodes are at a minimum (negative) voltage, thus trapping the conduction electrons in the pixel in which they are produced. Due to losses within the system only a fraction of the absorbed photons produce electron-

hole pairs and are trapped, this fraction is referred to as the quantum efficiency of the device. Channel stops, or insulating regions between columns prevent the charge from spreading across the columns. During a stare mode exposure the telescope is physically tracked such that light from a given object is incident on the same collection of pixels throughout the exposure. The "image" thus acquired by the CCD is a two-dimensional spatial distribution of charge where the amount of charge at any point on the CCD is proportional to the amount of light incident on the focal plane at that location.

At the end of the exposure rows of the accumulated electrons are "clocked" in parallel along columns of the CCD to the serial shift register by means of systematic adjustments of electrode voltage (parallel clocking), ordered to move the charge without changing the spatial distribution. To move the charge collected by all the pixels from their B electrodes to the C electrode (Figure 24), the voltage in all of the C electrodes is raised to its maximum value without changing the voltage in the B electrodes, such that the charge in each pixel can distribute itself between the two electrodes. Reducing the voltage of electrode B forces the charge to shift entirely into electrode C. Raising the voltage of the A electrodes and subsequently reducing that of the C electrodes, followed by the same procedure with electrodes A and B, will shift the entire array of charge by one pixel. The first (rightmost in Figure 24) row of charge on the CCD will now be in the central electrode of an orthogonal shift register (which has the same three-electrode structure as the body of the CCD, but oriented perpendicularly) where it is clocked in the same fashion as above out to a read-out device which records the charge and the order in which it is received. Once the orthogonal register is read, the parallel registers are clocked to

shift the entire array by another pixel and the second row is read off of the orthogonal register. The process repeats shifting the entire array by one row at a time until the entire CCD has been read out (see Kitchin 2003 for more details).
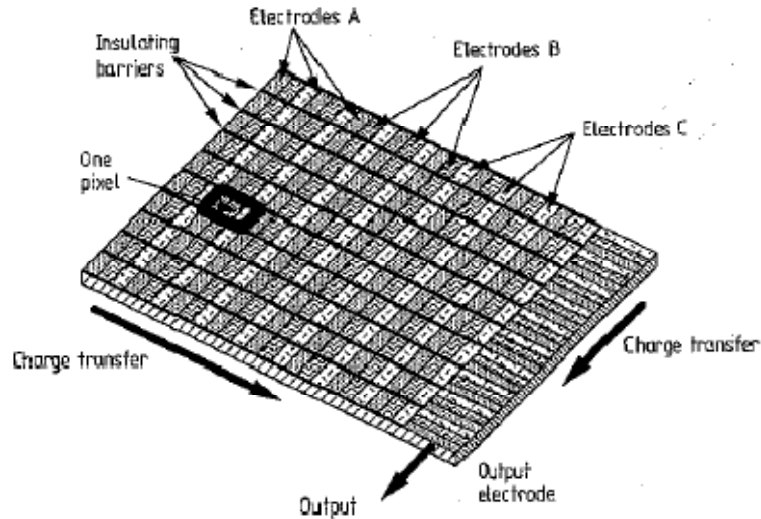
Figure 24. Three-phase CCD schematic. The parallel shift register runs along the columns (towards the lower right) and the orthogonal shift register is the right-most collection of electrodes which moves charge to the output electrode (from Kitchin 2003).

TDI mode differs from standard stare-mode CCD operation in a number of key ways. Because a TDI telescope is not tracked at the sidereal rate, photons from a given celestial object do not remain incident on the same pixel for the entire exposure, but instead the stellar images drift across the CCD at the sidereal rate. Under the operation described in the previous paragraph, this would result in an image of star trails. TDI mode counters this by shifting the electrons across the CCD in the same manner as for read-out, but does so continuously throughout the exposure, such that the accumulating charge is shifted from pixel to pixel in the direction of sidereal motion at the sidereal rate. Rather than having a lengthy read-out time at the end of an exposure, the TDI image is constantly being read-out with no dead time. Each star is exposed for the length of time that it takes

60

to transit the field of view of the CCD, while the entire night's observation could conceivably be composed of a single very long image or strip.

The image of the sky on a flat focal plane (ignoring any distortions due to telescope optics) is a gnomonic, or tangent plane projection of the celestial sphere. Great circles such as celestial meridians (lines of longitude) and the equator are projected as straight lines (the meridians radiate away from the poles), while small circles are projected as curved arcs (or complete circles around the pole). Stars on the celestial equator move in straight lines with the rotation of the Earth, while the projected paths have increasing radii of curvature with distance from the equator (as illustrated by the star trails in Figure 25). The path of a non-equatorial star across a CCD is a shallow arc parceled from a small circle centered on Earth's rotational axis (Figure 26). At any point in time, the position of a star in radians relative to the center of the focal plane is found using (Stone *et al.* 1996):

$$\xi = -\frac{\cos\delta \sin h}{\sin\delta \sin\delta_0 + \cos\delta \cos\delta_0 \cos h} \tag{25}$$

$$\eta = \frac{\sin\delta \cos\delta_0 - \cos\delta \sin\delta_0 \cos h}{\sin\delta \sin\delta_0 + \cos\delta \cos\delta_0 \cos h}, \tag{26}$$

where ξ and η are the positions along the east-west and north-south axes respectively, δ_0 is the declination of the center of the focal plane, δ is the declination of the star and h is the hour angle of the star. For a CCD inclined to the north-south axis by some angle i, the pixel positions, (x, y), relative to the center of the focal plane, (x_0, y_0), of the same star are (Stone *et al.* 1996),

61

$$x = x_0 + \frac{206264.806(\xi \cos i + \eta \sin i)}{s_x} \tag{27}$$

$$y = y_0 + \frac{206264.806(-\xi \sin i + \eta \cos i)}{s_y}, \tag{28}$$

where s_x and s_y are the row and column plate scales respectively. If the CCD is aligned accurately with the celestial meridian, these positions become functions of the east-west coordinate (Equation (25)) only.

Figure 25. Star trails over Mauna Kea illustrating the dependence of the projected paths of stars on the sky (or a focal plane). Note that the length of the trails and radii of curvature decrease towards the poles. Photo by Michael Michaud, Gemini Observatory/AURA.

In TDI mode using a stationary telescope, the image of a star is the gnomonic projection of the star's path integrated across the width of the CCD. If the CCD is aligned with the cardinal directions (north-south CCD axis on the meridian) and clocked at the sidereal

rate, the image of a star will be elongated north-south by the depth of the curved path. A non-sidereal tracking rate will elongate the path in the east-west direction. The image of a star suffering both a large curvature of path and incorrect clocking rate will be kidney bean-shaped with the concave side pointed north (or south in the southern hemisphere). Rotating the CCD with respect to the meridian will increase both of these effects. Observing on or near the equator combined with using a well aligned narrow CCD will reduce the elongation in declination while using an appropriate sidereal rate will reduce the elongation in right ascension.
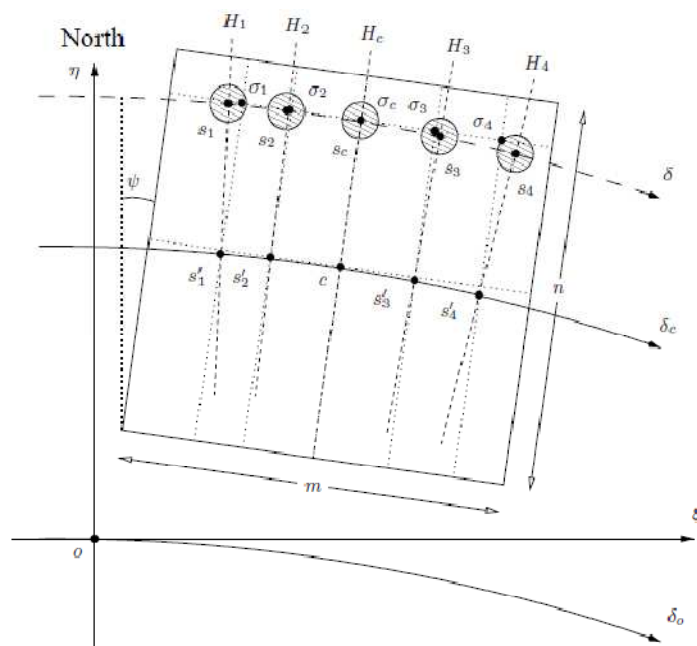
Figure 26. Projection of stellar paths onto a focal plane. The dashed lines labeled H_i indicate the celestial meridians at several hour angles subtended by the field of view while dotted lines indicate CCD rows and columns. The clocking rate is matched to the central declination. The positions of a studied star and the pixel aligned with the star's image at the meridian are indicated as a function of time by s_i and σ_i respectively. (From Vangeyte *et al.* 2002.)

The finite height of a CCD detector creates additional complications. It is clear from Figure 25 that the projected paths of stars near the poles are shorter than the projected paths of stars near the equator for the same exposure time, *i.e.* the gnomonic projection of

63

the sidereal rate varies as a function of declination. For a CCD clocked at a uniform rate matched to the sidereal rate at the center of the FOV, stars crossing the focal plane north of the centerline will lag the clocking rate, while stars south of the centerline will move faster than the clocking rate. The further a star is in declination from the clocking declination, the more severe will be the resulting image spread in RA. Because the projected sidereal rate is a function of the cosine of the declination, the greater the declination observed, the more severe the spread in the rates of the stars crossing a given field of view will be. The path of an off-center star and the ideal path that would produce a perfectly round image during an exposure are illustrated in Figure 26, clearly demonstrating both the effects of curvature and non-ideal clocking rate.

To demonstrate the specific TDI distortions we might expect in an image we can use as an example the 1.0 meter telescope at the US Naval Observatory. During the Flagstaff observing run (see Section 4.1), the CCD used at the 1.0 meter had a 13 arcminute field of view in declination and clocking rate of 30300 microseconds per row matched to the centerline declination of +35.2 degrees. The CCD was aligned with the meridian. Stars offset by six minutes of Dec will drift across the focal plane at rates offset from the clocking rate by 43 microseconds per row. Over the course of the exposure (with the CCD 4096 rows wide) the accumulated RA spread will be nearly six pixels, or 2.2 arcseconds using the plate scale of 0.38 arcseconds per pixel (top plot of Figure 27). The typical seeing at NOFS is about two arcseconds. The elongation in declination can be calculated from Equations (29)(25) through (28) by determining the difference in y-position of the same stars on entering and leaving the CCD relative to the y-position at

the center of the CCD. The CCD covers a range in hour angle of 0°.43 with the center of

the chip at an hour angle of zero. The resulting distortion in the y-direction is found to be

slightly less than two pixels or 0.716 arcseconds, with trivial variation across the field of

view at the declination observed (bottom plot of Figure 27). Assuming the CCD is

precisely aligned with the meridian, these distortions are not expected to adversely affect

the astrometric accuracy of our data. The consistency of the y-elongation to within a few

milliarcseconds means that any offsets in y-position due to these distortions will be

constant within our centroiding precision across the field of view, and more importantly,

as a function of time. The x-distortions are symmetric about the meridian crossing x-

position and will not change the x-centroid of the star.
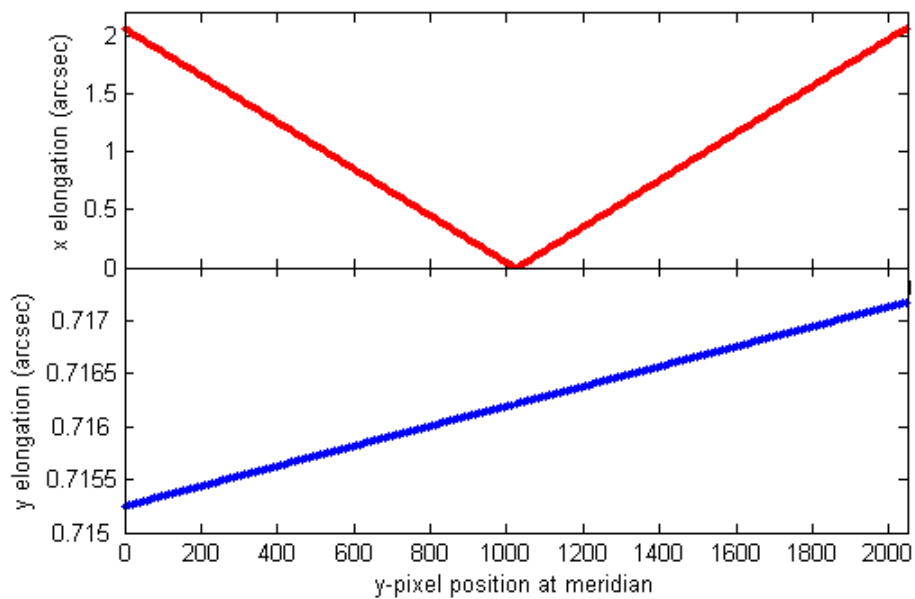
Figure 27. Elongation of a point source as a function of y-pixel position. The CCD clocking rate is matched to the declination of the chip center with increasing inconsistency between the clocking rate and sidereal rate with distance from the chip center causing the elongation in the x-direction. Curvature of the stellar paths causes the y-elongation.

Once the issues of CCD size, alignment and clocking rate have been negotiated, TDI

mode is an ideal platform for a variety of astronomical pursuits. The non-tracked nature

of observation means that telescopes can be simple in design without complicated tracking systems and more time and money can be put into quality optics and structural invariance. The ability to continuously image great swaths of the sky with no dead time for read-out makes TDI the method of choice for surveys seeking to efficiently image large parts of the sky in the smallest time frame.

CTI operated in the simplest possible manner, imaging the sky at +28 degrees declination continuously. Over the course of a year CTI observed a complete circle on the celestial sphere with nightly observations of each right ascension for up to several months at a time. Nightly observations on CTI consisted of several consecutive TDI "sweeps", each covering approximately 1.6 hours in right ascension and several minutes in declination. This mode of operation made CTI ideal for not only astrometric measurements (due to the numerous repeated observations of each object), but also for studying variable objects (binaries, supernovae) and moving objects (*e.g.* comets, asteroids and high proper motion stars).

The CTI telescope (see Benedict *el al.* 1991), which was operated on Kitt Peak between 1985 and 1992, employed a 1.8 meter f/2.2 Paul-Baker optical system with a focal plane populated by two RCA 320x512 30 micron pixel CCDs with scales of 1.5 arcseconds per pixel. The telescope structure was designed to be highly temperature invariant, minimizing nightly telescope focus changes.

For the analysis of the CTI data we chose several datasets of between 8 and 10 sweeps with as much overlap as possible. Positions in each sweep were centroided using Source Extractor windowed positions (Bertin & Arnout 2002, see Chapter 4 for further details). Positions of stars located in all sweeps were averaged together and the resulting set of positions constituted a secondary standard catalog. The number of sweeps applied to this averaging procedure was necessarily small due primarily to a combination of our requirement for continuous data of at least a few tens of minutes length and the difficulty of finding any number of science-quality sweeps with significant overlap. Due to the relatively small number of positions used in the averaging, significant errors were inherent in the secondary standard catalog; however, we expect the short timescale anomalous refraction to be variable enough on a night to night basis that its signature should be apparent when comparing a single night to the averaged data. This analysis should allow us to observe the minute to tens of minutes timescale anomalous refraction that is variable on a night to night basis and obtain a general idea of the timescales and amplitudes of the effect.

The standard positions were subtracted from corresponding star positions in each night's positional data to produce the residuals. Residual motions due to long timescale drift of the telescope were removed from the data by fitting a quadratic curve to the data and subtracting it from the residuals. Although the telescope was designed to be stable, the instrument still experienced structural distortion at the arcsecond level over the course of the night due to temperature changes. An example of CTI residuals is plotted in Figure 28.
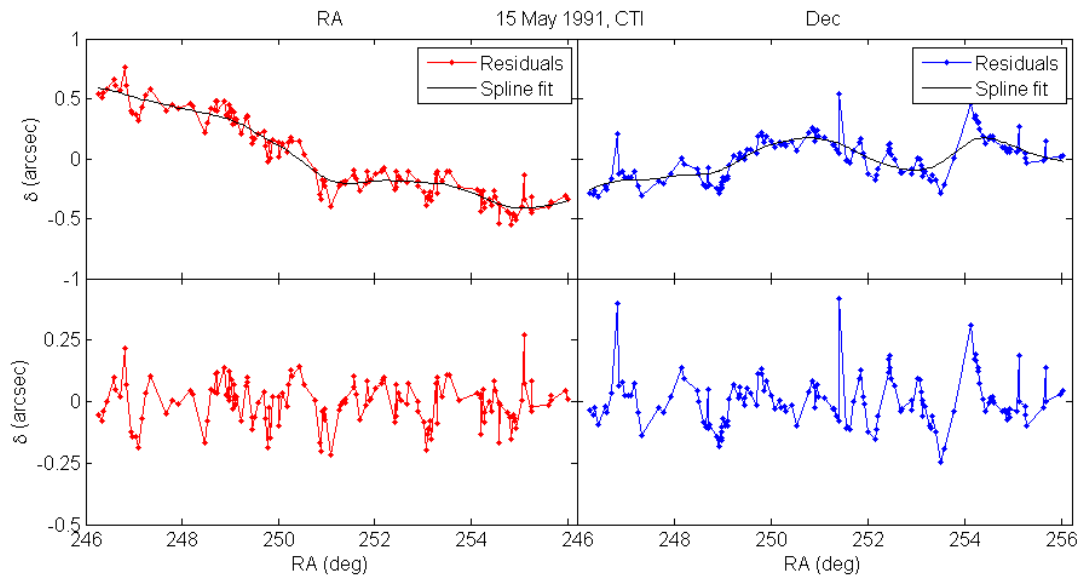
Figure 28. Example of CTI residuals. Each point is the difference between the position of a star on the night in question and the mean position of the same star over eight nights. The 500 brightest objects in each of the overlapping 1.6 hour sweeps are pared down to those stars which are contained in all eight sweeps. Top: RA and Dec residuals and spline fits to the longer period residual trends. Bottom: RA and Dec residuals with the spline fits removed highlighting the shorter period residuals. Fifteen degrees of RA correspond to one hour.

Initial examination of several nights' worth of CTI data shows variations in the residuals on several scales. Long timescale variations have amplitudes ranging from a few tenths of an arcsecond to as much as an arcsecond (although the errors inherent in the secondary standard catalog likely contribute to the amplitudes of these oscillations) and generally have periods of several tens of minutes. The smaller scale oscillations typically have amplitudes of several tenths of an arcsecond and periods of a few minutes. Another example of CTI residuals and the associated power spectra is shown in Figure 29.

The CTI residuals shown in the Figure 28 and Figure 29 are highly representative of the residuals found in all CTI data. This examination of CTI astrometry lends further

evidence to the hypothesis that anomalous refraction is universal to all ground-based astrometric observations. We also see that AR appears to be a continuously occurring phenomenon; however, the nature of the CTI data and the residual finding method used could potentially induce the appearance of AR where it is not actually present in a particular night's positions.
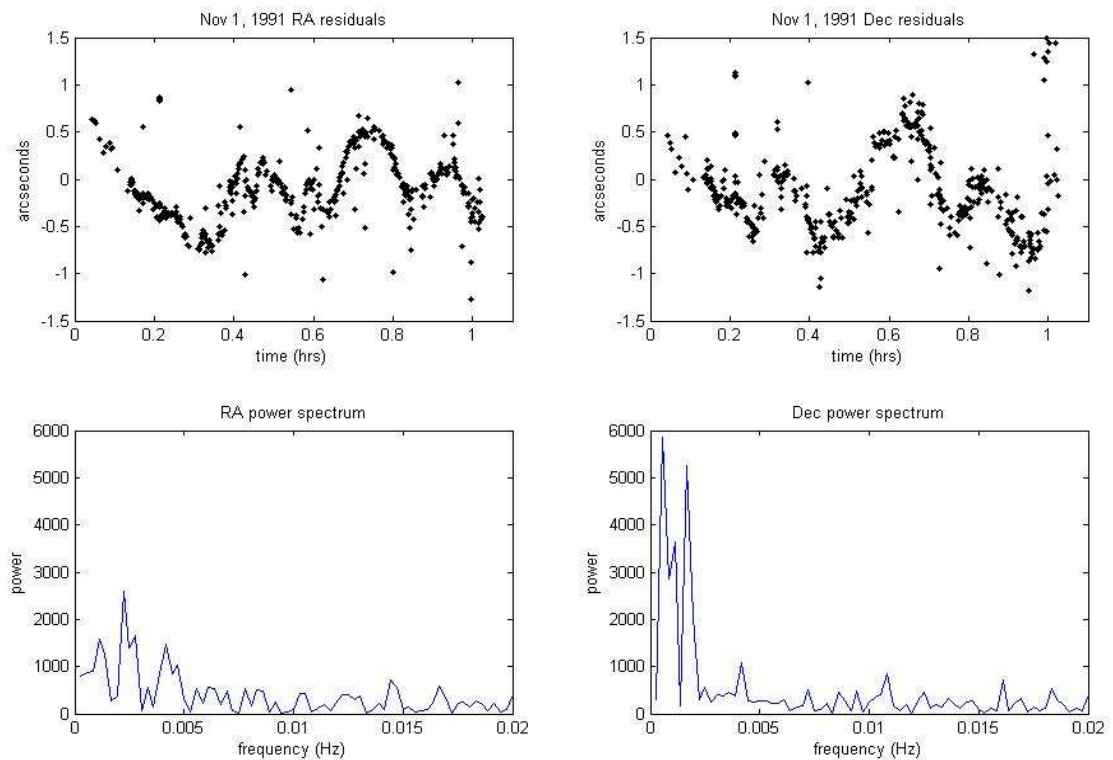
Figure 29. CTI residuals in RA and Dec as a function of time (RA) (top). Associated power spectra of RA and Dec residuals (bottom).

3.2 SDSS

The second set of astrometric data we are using for this analysis was from SDSS. Jeff Pier, the principal astrometry investigator on the Sloan survey during that telescope's commissioning phase, secured access to the unprocessed SDSS datasets for this research.

Pier's analysis of the early SDSS data resulted in one of the more comprehensive discussions of anomalous refraction to date (Pier *et al.* 2003). We reprocessed the raw datasets using the SDSS reduction pipeline, modified to remove long (2 hour+) period thermal and mechanical telescope motion, but not the anomalous refraction. Having access to this extraordinary dataset allowed an incredibly in-depth study of the characteristics of anomalous refraction affecting the Sloan telescope.

The Sloan telescope is a 2.5 meter modified Ritchey-Chretien system designed for very wide field of view observation. The telescope is operated at the Apache Point Observatory in the Sacramento Mountains of southern New Mexico. In order to minimize terrain-induced turbulence, the instrument is mounted on a platform that extends out over the windward side of the mountain, several meters above the surface and vegetation. The telescope is housed in a building that rolls completely away from the telescope during observation, leaving the instrument exposed to the ambient air and removing any effects of dome-seeing. A baffle surrounding the telescope structure (but mechanically isolated from the telescope) is designed to minimize wind buffeting and reduce stray light in the optical system.

The data archives of the SDSS include millions of images taken over the last nine years. These data have several qualities which make them ideal for this study of anomalous refraction. The Sloan camera has an edge-to-edge field of view of 2º.3 and a focal plane array consisting of 30 Scientific Imaging Technologies (SITe) photometric CCDs organized into five filter bandpasses and 22 (2048x400) astrometric CCDs (Figure 30).

Each photometric CCD has 2048x2048 24 micron pixels with a pixel scale of 0".396/pixel (Pier *et al.* 2003). The availability of color data allows analysis of dispersion effects, and the wide angle span of the focal plane makes potential observation of phase lag (*i.e.* the changing tilt of the atmosphere as waves pass overhead) possible.
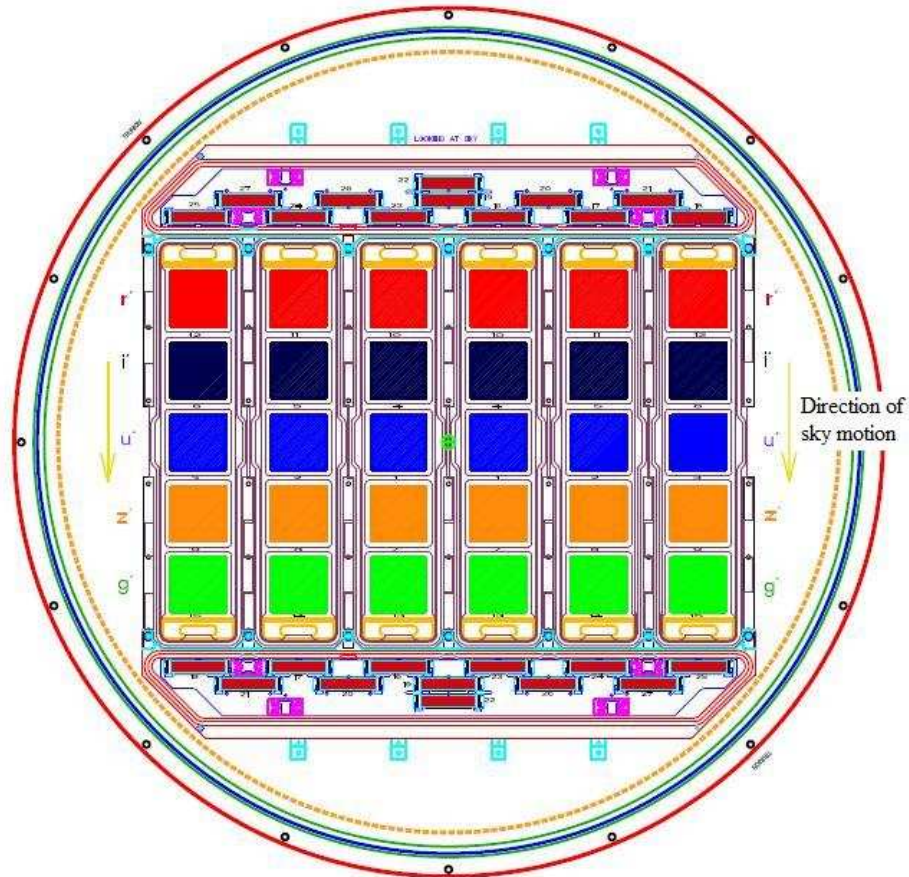
Figure 30. SDSS focal plane array (from Pier *et al.* 2003). The imaging camera has 30 photometry CCDs arranged in five filters (r', i', u', z' g'). Small rectangular CCDs at the front and end of the array were originally used for astrometry.

Like CTI, SDSS uses the time-delay and integrate (TDI) readout mode. Unlike CTI, however, they do not employ a stationary telescope. SDSS observes strips following great circles in the SDSS survey coordinate system. Because SDSS is a survey designed to observe most of the sky, including high declinations, and because the camera has a

71

very wide field of view, the issue of differences in the celestial drift rate across the declination range of the FOV is not to be taken lightly. By observing along great circles that intersect the local meridian at right angles instead of lines of constant declination, SDSS uses the fact that great circles on the sky are projected as straight lines in the gnomonic projection to avoid TDI distortion, allowing observation of high declinations without compromising image quality. The telescope is tracked along these great circles such that the time for objects to cross the field of view (54 seconds) remains uniform regardless of the altitude or direction at which the telescope is tracking. The sky drifts across the telescope along CCD columns, from the *r'* CCDs to the *g'* CCDs (see Figure 30). There are 72 seconds between when a given star crosses the first row of one CCD and when it crosses the first row of the next CCD in the column for a total of 288 seconds for the star to cross the full field of view.

The survey coordinate system is divided up into 2.5 degree separated great circles each defined by a stripe number, where stripes 10 and 82 are located at the equator. Because the columns of CCDs along the scanning direction are separated by 80% of the CCD width, each stripe is actually composed of a north and south strip, observed on different nights and offset by almost a CCD width to allow full coverage of the given stripe.

The survey employs the great circle coordinate system where latitude and longitude in a particular great circle strip are designated μ and ν, respectively, such that $\nu = 0$ along the center of the strip and μ increases along the scanning direction (Pier *et al*. 2003). On the equator stripes, μ and ν correspond approximately to right ascension and declination,

respectively, in the equatorial coordinate system. Images taken on the equator stripes employ the least amount of telescope tracking while the degree to which the telescope must be tracked increases with increased survey latitude.

3.2.1 Data Reduction

The astrometric reduction of the SDSS images involves an extensive data processing pipeline (*e.g.* Pier *et al.* 2003; York *et al.* 2000; SDSS Project Book 1999). Data are read off of the photometric CCDs in real time and divided into "frames" of 2048 columns by 1361 rows for processing purposes. Bright stars in the astrometric CCDs are centroided and based on these positions, stars in the photometric CCDs are located and centroided. The centroids are found by fitting a two-dimensional Gaussian to the stellar point-spread function (PSF). Variation in the PSFs as a function of position on the frame is determined and corrections are made to the centroids to remove any biases due to an asymmetric background. The *r'* CCDs are traditionally used as the astrometric reference for the all other CCDs (for our research all photometric CCDs were directly reduced against the standard catalog). Prior to the release of the US Naval Observatory CCD Astrograph Catalog (UCAC) catalog SDSS planned to use the smaller astrometric chips for matching against Tycho-2 catalog stars (which saturate in the photometric CCDs). For matching against UCAC these CCDs are no longer necessary because the faint magnitudes of stars included in the UCAC catalog are well matched to the magnitude range of the photometric CCDs.

The astrometric pipeline (*astrom*, for more information see Pier *et al.* 2003, Hennessy *et al.* 2006) takes the lists of positions, telescope metadata (pointing, timing, camera

rotation, *etc.*), weather conditions and a catalog of astrometric standard stars and outputs transformation coefficients which convert frame *x* and *y* coordinates to catalog mean place (CMP) coordinates on the sky. CMP is the position (of a celestial object) as seen from the solar system barycenter referenced to mean equator and equinox of the J2000 epoch. According to Pier *et al.* (2003), the minimum accuracy requirement for the SDSS astrometry was 180 milliarcseconds (mas) rms (root-mean squared) per coordinate to allow accurate placement of the survey spectroscopy fibers; however, excluding systematic errors, the final published SDSS astrometric accuracy is actually 45 mas rms.

The first step of the pipeline is, for each observed star, to apply x and y frame positions, position of the CCD relative to the boresight (the pointing axis of the camera), scale factors and distortions for both the telescope and camera, and camera rotation to the known focal plane model to determine the stellar position relative to the boresight CMP position (pseudo-catalog place, PCP). Specifically, pseudo-catalog place is defined as $PCP_* = CMP_{bs} + (OP_* - OP_{bs})$, where * and bs denote star and boresight positions respectively and OP is the observed place – the position as observed by an observer on the surface of the Earth. Pixel positions are corrected for third order focal plane distortions, which are only a function of column position due to the drift-scan observing (over an exposure a star will "see" all focal plane distortions that are a function of row, making them uniform for all stars). Row position (*x*) is converted to a time (when the star was mid-exposure) as, $t = C \times (x - x_{ref})$, where C is the clocking rate of the CCD (seconds of time per row), *x* is the row position of the star (row numbers increase continuously regardless of frame divisions starting from 0 at the start of the scan) and

x_{ref} is the readout row when the star was at mid-exposure. At time, t, during a scan the position of the boresight along the great circle (see Figure 31) is determined from a given star's position (assuming small angles) as:

$$\mu_{bore}(t) = \mu_0 + CT\big(x(t) - x_{ref}\big) \tag{29}$$

$$\nu_{bore}(t) = \nu_0 + \phi(t)\big(x(t) - x_{ref}\big), \tag{30}$$

where CT is the tracking rate in arcseconds per row, ϕ is the error angle in the boresight tracking relative to the desired tracking direction μ and μ_0 and ν_0 are the great circle CMP longitude and latitude at time, $t = 0$. The PCP position of a star is then,

$$
\begin{aligned}
\mu(t) = \mu_{bore}(t) + s\big[x_i - x_{bore} + (y_i - y_{bore})\theta_f \\
+ f_i\{x_c + y_c(\theta_f + \theta_i) + p_c f_c(y(t) - 1024)(\theta_f + \theta_i + \theta_c)\}\big]
\end{aligned}
\tag{31}
$$

$$
\begin{aligned}
\nu(t) = \nu_{bore}(t) + s\big[-(x_i - x_{bore})\theta_f + y_i - y_{bore} \\
+ f_i\{-x_c(\theta_f + \theta_i) + y_c + p_c f_c(y(t) - 1024)\}\big].
\end{aligned}
\tag{32}
$$

The telescope scale, s, is approximately 16.6 arcseconds per micron (SDSS Project Book 1999), while the scale correction factors of the Dewar relative to the telescope, f_i, and the CCD relative to the dewar, f_c, are both nominally 1.0. The pixel size, p_c, is 0.024 mm. The position of the reference pixel, $(x_{ref}, 1024)$, relative to the boresight is the position of the reference pixel relative to the dewar hardpoint, (x_c, y_c), plus the distance from the instrument rotator axis to the hardpoint, (x_i, y_i), minus the distance from the rotator axis to the boresight, (x_{bore}, y_{bore}). The angles θ_f, θ_i and θ_c correspond to the rotations of the instrument rotator relative to the tracking path, the dewar about its hardpoint and the CCD about its reference pixel respectively.

The conversion from great circle latitude and longitude to J2000 RA and Dec uses

$$\tan(\alpha - \mu_0) = \frac{\sin(\mu - \mu_0)\cos\nu\cos i - \sin\nu\sin i}{\cos(\mu - \mu_0)\cos\nu} \tag{33}$$

$$\sin\delta = \sin(\mu - \mu_0)\cos\nu\sin i + \sin\nu\cos i, \tag{34}$$

where α and δ are RA and Dec in J2000 coordinates, and i and μ_0 are the inclination and J2000 RA of the great circle ascending node, respectively (Pier *et al.* 2003).

Once the PCP for each frame is determined, the next step is to match the frames to a standard astrometric catalog (UCAC) using the positions found with the focal plane model as a starting point. Because the PCP positions use the great circle coordinate system, we must first convert the UCAC J2000 RA and Dec to μ and ν. The transformation from J2000 CMP to great circle PCP involves shifting barycentric UCAC positions to apparent topocentric positions and requires consideration of time scales, precession, nutation, aberration, polar motion and normal refraction (for more details see Kovalevsky & Seidelman 2004; Cox 2000). The Earth is unfortunately not a stable observing platform. We will consider each of these factors in turn.
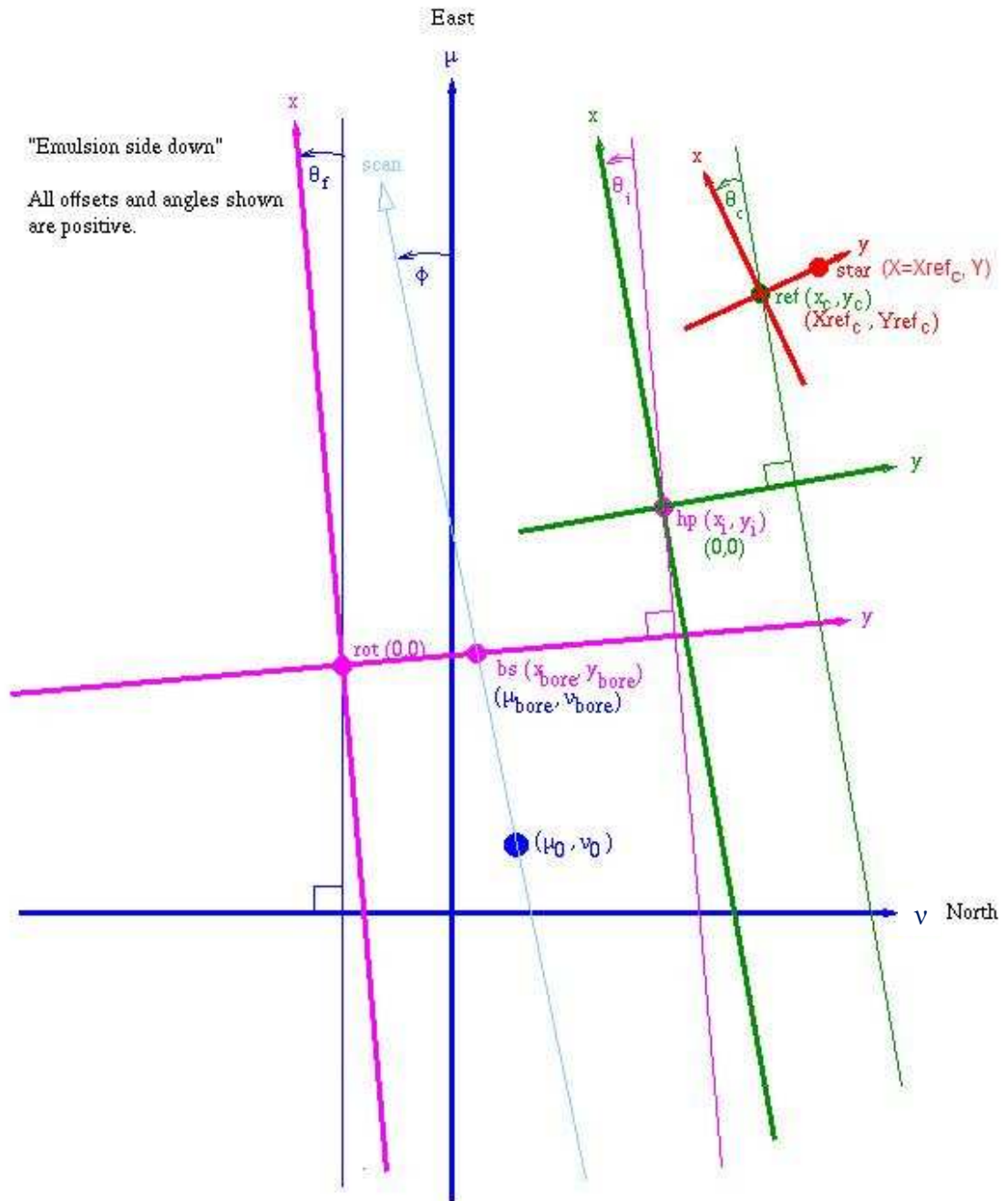
Figure 31. Focal plane model. μ and ν are the strip longitude and latitude (blue, desired tracking path is along μ). The instrument rotator axis (pink) is rotated by θ_f from μ and the boresight (bs, with strip position at time, t of (μ_{bore}, ν_{bore})) is offset from the rotator axis by (x_{bore}, y_{bore}) and has an instantaneous tracking path (light blue) angle error of φ. The dewar hardpoint (hp, green) is offset from the instrument rotator axis by (x_i, y_i) and has a rotation relative to the rotator axis of θ_i. The CCD reference pixel (red, with pixel position $(Xref_c, Yref_c)$) is offset from the dewar hardpoint by (x_c, y_c) and is rotated by θ_c relative to the dewar. The position of a star is defined at the reference row as $(X = Xref_c, Y)$. Adapted from Hennessy *et al.* 2006.

The standard time system familiar to most astronomers is Universal Time (UT) which is derived from solar time and corrected for polar motion (also UT1). International Atomic Time (TAI) is based on the SI second defined by vibration of the Cesium-133 atom. For terrestrial astrometry the Terrestrial Dynamical Time (TDT) is the idealized time on the geoid of the Earth (surface corresponding to mean sea level) and is approximately TAI + 32.184 seconds. Barycentric coordinate systems employ Barycentric Dynamical Time (TDB) which includes relativistic corrections for referring equations of motion to the solar system barycenter. TDB varies periodically with respect to TDT with a maximum difference between the timescales of 0.001625 seconds. Because this corresponds to a change in RA of only 24 mas and varies on timescales of years not minutes, only TDT epochs are used.

Corrections for precession are an absolutely vital part of the transformation between coordinate systems and epochs. Precession is the long period motion due to the combined torques of the sun, moon and planets on the Earth which cause the rotational axis of the Earth to wobble about the pole of the ecliptic with a period of 25,800 years. The precession in longitude is 5029".0966 per Julian century or roughly 138 mas per day. This is significant considering that the anomalous refraction effects we are studying have amplitudes of hundreds of milliarcseconds, although the timescales of the two effects are different. Precession is a well known phenomenon (*e.g.* Kovalevsky & Seidelmann 2004; Lieske *et al.* 1977) and the actual derivations will not be discussed here.

The small amplitude short period oscillations of the Earth's axis due to the gravitational influences of other solar-system bodies are called nutation. The largest component of nutation has an amplitude of 9".0 and a period of 18.6 years, or a little less than half an arcsecond per year. Precession and nutation are often treated as a single effect (having the same source). Like precession, nutation is well studied (*e.g.* Kovalevsky & Seidelmann 2004) and while applied as part of the astrometric reductions, will not be addressed in this writing.

Aberration is caused by the combination of the finite speed of light and the motion of the observer's reference frame. Light travelling from a distant star encounters a telescope on the Earth with a velocity vector that is entirely radial with respect to the telescope; however, because the telescope is located in a reference frame that is moving with respect to the star, the vector of the telescope's motion must be added to the light's true velocity vector to get the apparent velocity vector. A good analogy is a north wind, that when observed from a bicycle travelling towards the east, appears to be incident from the north-east. There are three components to the motion of our telescope that need to be considered when solving for the effects of aberration. The daily motion due to the Earth's rotation causes diurnal aberration, the annual motion due to the Earth's orbit results in annual aberration and that due to the solar system's motion through space gives rise to secular aberration. Quantitatively the aberration is approximately $V/c \sin\theta$, where V is the magnitude of the observer's motion, c is the speed of light and θ is the angle between the two velocity vectors. The diurnal aberration is $0.320 \sin\theta_D$ arcseconds, where θ_D changes by 15 degrees per hour. The annual

aberration is approximately $20.5 \sin \theta_A$ arcseconds, where θ_A changes by about one degree per day. For the purposes of this research secular aberration is constant and because the reference frame in which the barycenter of the solar system moves is ill-defined, there is little meaning attached to the absolute value of the secular aberration. In general the timescales of aberration are too long to significantly affect our anomalous refraction analysis; however the annual and diurnal components are accounted for in our data reduction.

The above mentioned effects are applied in the conversion from barycentric (solar-system center of mass centered) place to topocentric (Earth surface centered) place. To achieve positions in observed place we need to also account for the atmospheric refraction due to zenith angle, which was discussed in the previous chapter and is defined in Equation (23).

The frames are matched to the astrometric standard catalog (UCAC) based on the calculated positions using a linear least-squares method fitting the boresight tracking terms (Pier *et al.* 2003). A piece-wise polynomial smoothing spline fit is applied to both the stripe latitude and longitude residuals separately as a function of time (with knots placed every 5-20 frames under standard operation) to remove remaining systematics. For the purposes of this study, we applied a smoothing spline fit to the residuals with knots placed every 200 frames (two hours) to remove telescope motion, while leaving any anomalous refraction effects in the data.

The output of the SDSS pipeline includes lists of the mean and median offsets from UCAC of each frame in a stripe and for each CCD, as well as the frame times, quality indicators such as number of stars per frame, telescope metadata such as tracking and camera temperature, and weather conditions during the observation.

3.2.2 SDSS Results

The majority of the data analysis in this research focused on stripe 82 (equator stripe) because this stripe has the largest number of repeated runs (a run being a single night's observation of a given strip), and the least amount of telescope tracking. (Early engineering runs were taken on the equatorial stripe with no tracking as well.) This allows us to examine images taken on different nights and under varying weather conditions, but with the same telescope setup and operation. We chose seven runs based on data quality and run length, with all data taken between late September and early November of 2003, 2004 and 2005 (the limited seasonal range was not intentional). We also examined data from other stripes to compare residuals from runs where the telescope was significantly tracked, or data taken at varying zenith angles. The stripes used for this data were 22 and 24 with maximum declinations of 30 and 35 degrees respectively (low zenith angle), stripe 39 with maximum Dec of 75 degrees, stripe 44 at 84 degrees and stripe 79 at 8 degrees. The stripes with the highest declinations involved the most telescope tracking.

The first order in the study of the SDSS residuals was to confirm the existence of anomalous refraction in the data and to ascertain whether the effect was continuously occurring or intermittent. The residuals from all seven stripe 82 runs analyzed are shown

in Figure 32 as a function of stripe longitude (approximately RA for equatorial stripes). 15 degrees correspond to one hour (also 100 frames). Each point on the image is the mean offset of all stars in a single frame from their UCAC catalog positions. Anomalous refraction occurs continuously and on all nights with similar amplitudes (tenths of arcseconds) and periods (a few minutes to tens of minutes) in all data. The same continuity of AR is also seen in the data taken from non-equator stripes.
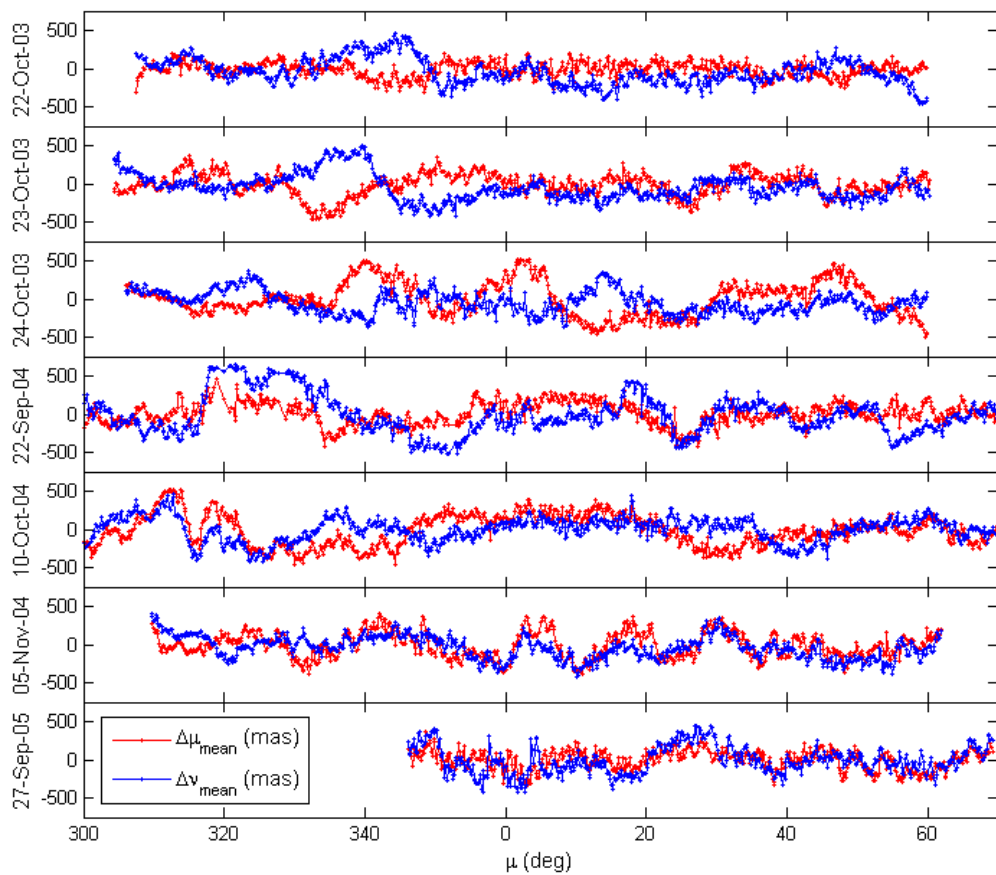
Figure 32. Residuals from all seven stripe 82 runs.

Figure 33 illustrates the frequency and amplitude characteristics of a particular stripe 82 run (4207, 24-Oct-03). Both high and low frequency residuals are clear in the figure with the low frequency residuals typically having larger amplitudes than the high frequency.

82

A Fourier spectrum indicates dominant periods of about five to 15 minutes for the high frequencies and several tens of minutes to hours for the low frequencies (most of the low frequency peaks have amplitudes truncated by the figure).
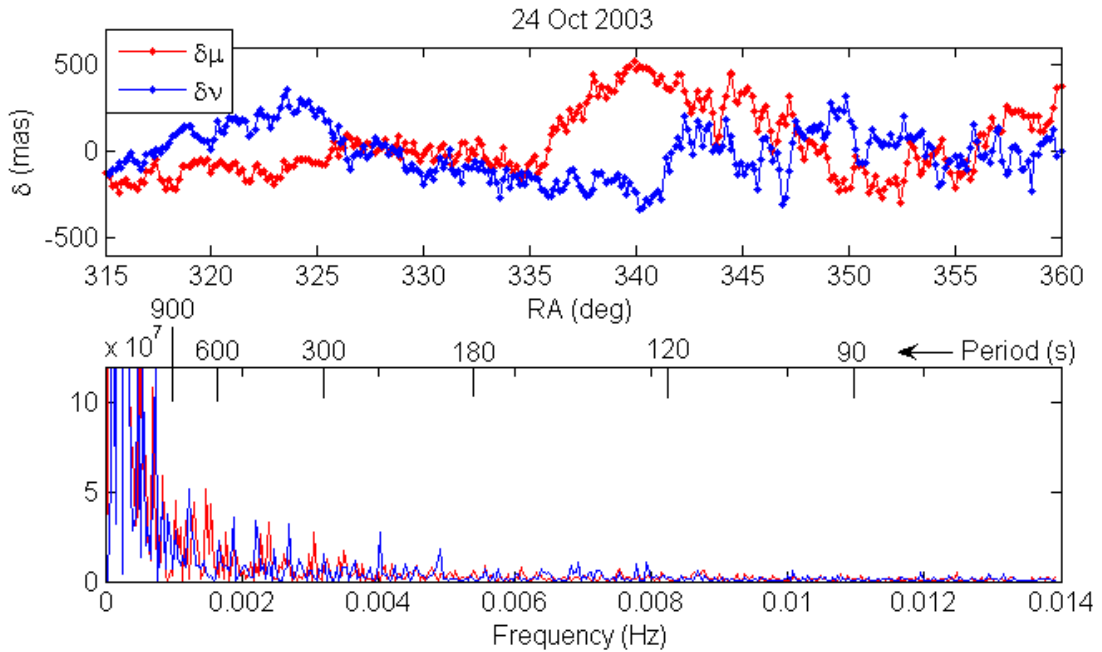
Figure 33. Segment (approx. 3 hours of time) of run 4207 short period (smoothing spline subtracted) mu (RA) and nu (Dec) residuals and frequency spectrum of the full run of spline-subtracted residuals.

Our next concern was to examine the coherence of AR across the 30 CCD, 2°.3 FOV of the SDSS camera. The coherence across the FOV is illustrated in Figure 34 and Figure 35. Figure 34 shows the residuals from all of the *r'* CCDs (across the first row of CCDs) for a two hour section of a single run, while Figure 35 shows the same thing, but for the first CCD of each filter (*i.e.* down the first column of CCDs). At first glance there appears to be a high degree of correlation across the field of view, particularly at the longest timescales. The level of correlation decreases with decreasing timescales of residual motions.
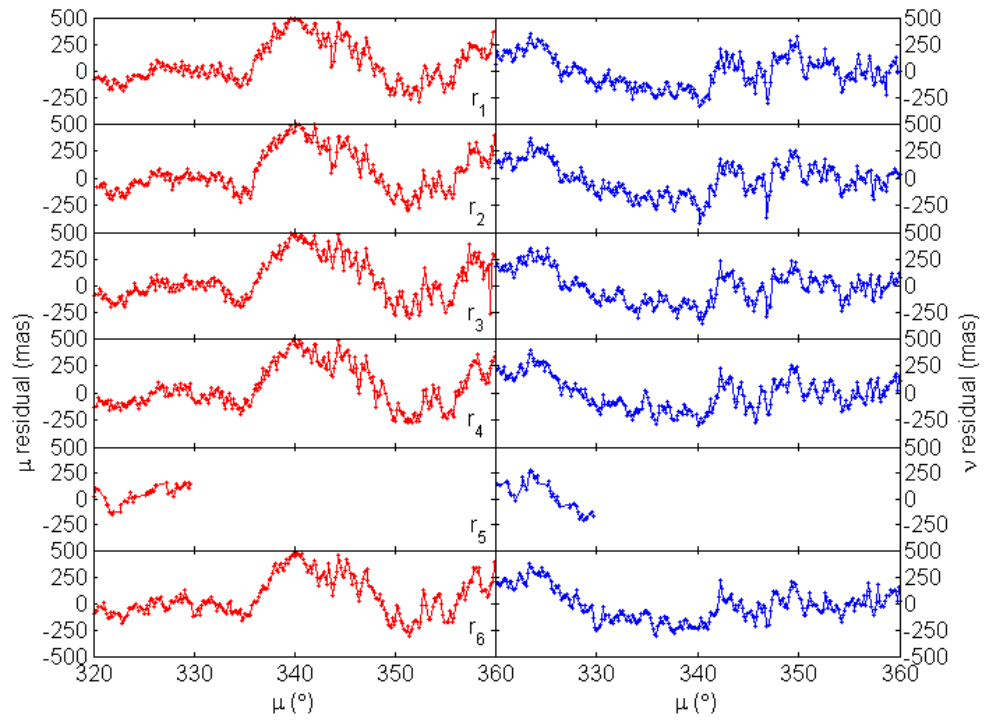
Figure 34. Two hour section of run 4207 residuals, all *r'* CCDs (across focal plane perpendicular to scan direction). CCD *r5* was malfunctioning during the run examined.
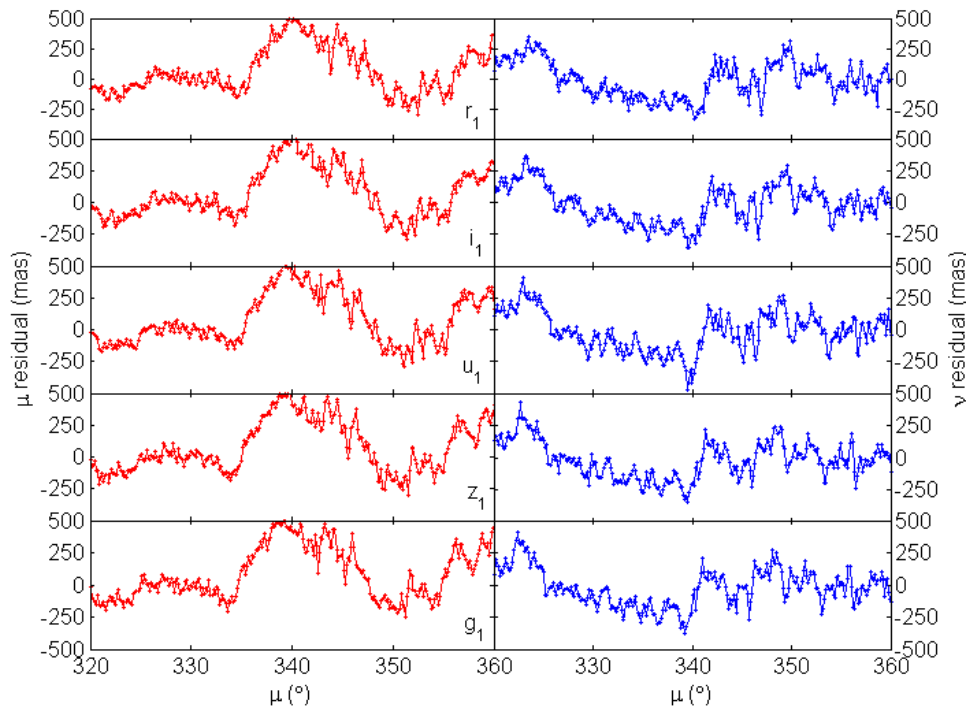
Figure 35. Same as **Figure 34**, but down the first column of CCDs (first CCD of all five filters).

84

Close examination of the μ and v residuals in Figure 35 suggests a significant phase lag; however, this is a result of a data processing step that offsets frame numbers such that each CCD sees the same stars in the same frame (otherwise a given star is imaged by each successive row of CCDs at a later frame number). This causes the times for when a specific frame was observed to be offset and because AR is time dependent, the frame offset appears as a phase lag. Cross-correlations of CCDs at the opposite ends of the field of view (Figure 39, also Figure 37 and Figure 38) give no indication of any phase lags present in the SDSS data.

The study of the SDSS cross-FOV correlation was of particular interest because this was our only dataset where multiple independent measures of anomalous refraction were available across a wide range of angles from a single instrument. Correlation across the CCDs indicates the minimum scale of anomalous refraction, but lack of correlation, or modulation of residuals across the FOV not only indicates that AR must be independent of the telescope, but also provides a measure of the scales over which anomalous refraction *changes*. This last consideration is of particular importance when considering the wide field astrometric accuracy of very large FOV instruments (*i.e.* LSST and panSTARRS). Both the coherence and modulation of anomalous refraction can be demonstrated by mapping the focal plane errors as a function of time. For each frame (approximately 36 seconds of time per frame) the mean AR offset in μ and v in each of the 28 (functional) CCDs was determined and converted into a vector. We can then create a simple movie showing the FOV offset and distortion as a function of time: Focal Plane Map Movie (see also Appendix I for screen shots of selected movie frames). The

arrows are anchored at the approximate relative positions of each CCD and the arrow length is scaled such that the distance between ticks on the graph corresponds to one arcsecond. The uniform motion of all of the vectors clearly indicates the correlation across all of the CCDs, while the more subtle relative motions between the arrows shows the focal plane distortion. The latter effect is highlighted by subtracting off the mean offset of the entire focal plane and looking at only the relative displacements of the positions in the individual CCDs for the same run: Focal Plane Distortion Movie (see also Appendix I). The distance between ticks on the graph in this case corresponds to 0.2 arcseconds. Finally, a three-dimensional surface plot is a nice illustration of the apparent tilt and distortion of a different run (note that all runs described here are from stripe 82): Focal Plane Tilt Movie (see also Appendix I). The tilt is the gradient (normal surface to the residual vector) of the focal plane at each frame. In this movie the CCDs are located at the corners of the grid.

Subtracting the residuals in selected frames from each other and plotting the resulting residual differences (Figure 36) highlights the departures from correlation. Both frames from adjacent CCDs and frames from CCDs located at opposite ends of the focal plane are compared in this manner. Subtracting the residuals in adjacent frames (top left) leaves mostly random high frequency noise with very little indication of systematic differences. This is also clear in the power spectrum (top right) of the residual differences. Performing the same subtraction with CCDs located at opposite ends of a column (second pair of plots), row (third pair) and diagonal (bottom set) indicates a notable increase in systematic differences between the residuals. Of particular interest

86

are the peaks at periods of order 10 minutes which appear in the power spectra of the residual differences between the opposite CCDs. This strongly suggests that the source of AR with periods of a few tens of minutes subtends angular scales comparable to the size of the SDSS field of view.
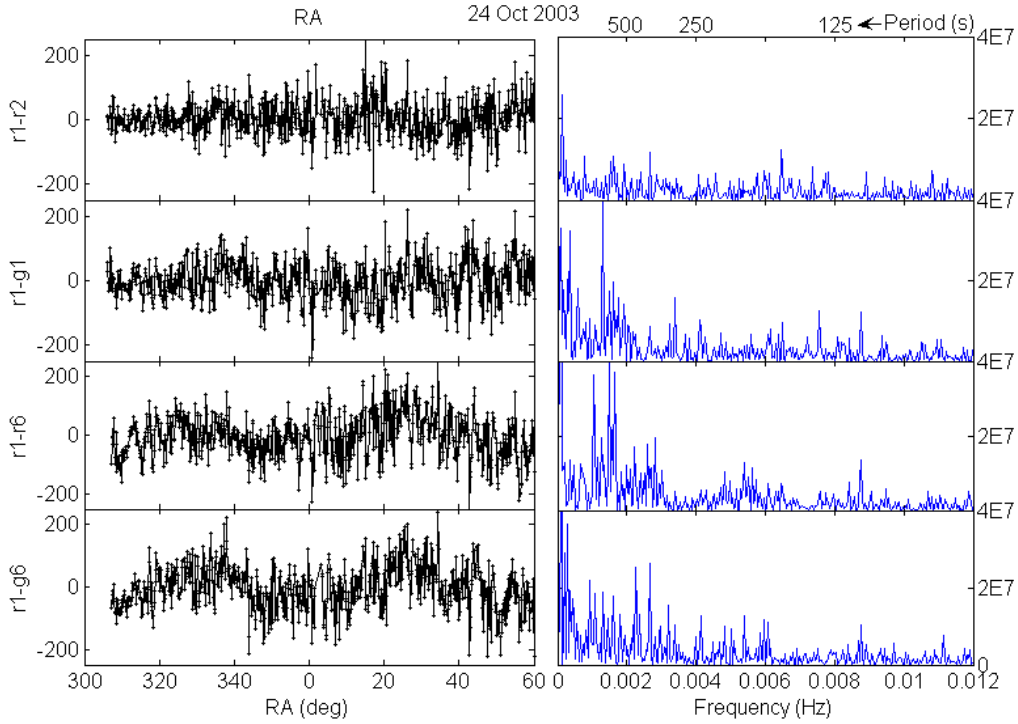
Figure 36. Selected comparisons between simultaneous residuals on different CCDs. Top row of plots is the difference between residuals on adjacent CCDs and associated power spectra. Second row shows the difference between CCDs at opposite ends of a column and associated power spectra. Third row shows the difference between CCDs at opposite ends of a row and associated power spectra. Last row shows difference between the CCD at upper left corner of focal plane and the CCD at lower right corner and associated power spectra.

We further examine the correlations across the FOV by comparing the residuals from CCD $r'1$ and CCD $r'6$ (opposite ends of the focal plane). Overlaying the residuals from the two CCDs on a particular run (4933) illustrates the high correlation of the long period motions (Figure 37). An unbiased cross-correlation confirms that the residuals are closely related and also illustrates the strong periodicity in both sets of residuals.

Subtracting a smoothing spline fit to the low frequency residuals and performing the same cross-correlation analysis (Figure 38) indicates that the correlation is much weaker at high frequencies, although still significant.
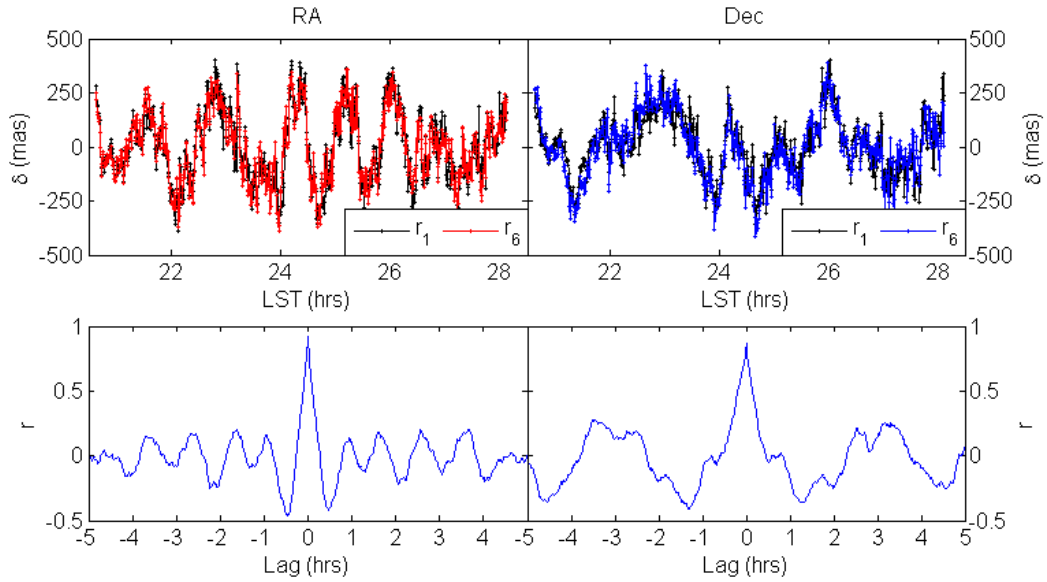
Figure 37. μ (RA) and ν (Dec) residuals from run 4933 (5-Nov-04) CCDs *r'*1 and *r'*6 overlaid to illustrate low frequency coherence (top) and the normalized cross-correlation coefficient, r (bottom).
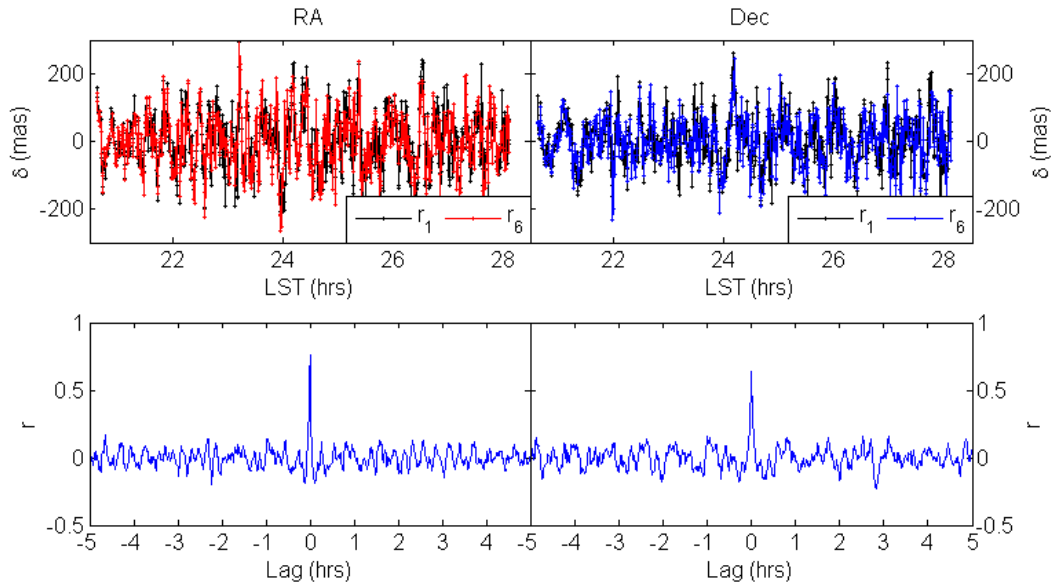
Figure 38. Same as above, but with a smoothing spline fit to the motions with of order hour timescales subtracted.

The unique focal plane CCD array of the Sloan telescope made this dataset our single source of potential dispersion data. Based on the atmospheric refractivity equations presented in the theory section (Equations (19) through (22)), we can determine the spread in refraction amplitudes we expect to measure across the wavelength range of the SDSS CCD array. The five Sloan filters have effective wavelengths of 354.9-nm, 477.4-nm, 623.1-nm, 761.5-nm and 913.2-nm (*u', g', r', i',* and *z'* respectively, Gunn *et al.* 1998). For a 482 nm wavelength difference (*u' – z'*) under standard atmospheric conditions we can expect a refraction amplitude difference of approximately 5%. For an average anomalous refraction amplitude of 0.5 arcseconds, this would translate into about a 0.025 arcsecond difference in residual amplitude between the two filters mentioned.

Clearly, optical wavelength dispersion observed in the anomalous refraction residuals would be a key indicator of the atmospheric refractive source of the effect. Unfortunately, intensive examination of the SDSS data shows no indication of dispersion. The root of this absence lies in the astrometric accuracy of the Sloan dataset. The original purpose of the SDSS astrometric data reduction was to provide positions of objects suitable for the placement of the fiber-optic cables used in the acquisition of spectra. This led to a published astrometric accuracy of 45 mas rms (Pier *et al.* 2003). It is probable to suggest that finding a systematic 25 mas difference in refraction in this dataset is an impossible task. Nevertheless, the attempt must be made with the caveat that only certain AR scales will result in the maximum dispersion of 25 mas.

Given the Sloan telescope field of view of $2°.3$, and assuming that residuals are caused by a wave at one kilometer above the telescope; in order to see the high level of correlation on all CCDs we require that the wavelength be greater than the linear field (approximately 40 m, see Figure 23) viewed by the telescope at that altitude. At time, $t = 0$, a simple monochromatic wave can be described by $y = A \cos kx$ and the slope of the wave by $y' = kA \sin kx$. If we set the position of one end of the field of view at one kilometer to zero, then the slope of the wave at that point is also zero and we require that at the other end $y'(40) = kA \sin 40k < \tan \alpha_c$ where α_c is the critical angle of tilt of the atmosphere above which we begin to see anomalous refraction. If we use the published astrometric accuracy of the SDSS as the smallest AR we can observe, then based on Figure 13, $\alpha_c \cong 13°$ and the ratio of the wave amplitude, A, to the wavelength, $\lambda = 2\pi/k$, at the critical angle is approximately 0.08. From this we conclude that the minimum wavelength for a wave at one kilometer to be correlated across all CCDs is 525 meters. This minimum wavelength is proportional to the altitude of the wave. For a wave occurring at an altitude of 250 meters the same reasoning gives a minimum wavelength of 132 meters.

If the wavelength of the wave is shorter than the critical value for a given altitude we can expect to see a phase lag between residuals observed on one end of the focal plane and those observed from the other end. If the wavelength is half of the critical value (but with the same amplitude as the critical wave), the angle of atmospheric tilt will change by almost 40 degrees across 40 meters corresponding to a difference of refraction of nearly two tenths of an arcsecond across the field of view. (*i.e.* the FOV of the telescope will

span 15% of the wavelength of a 270 meter wave at one kilometer above the telescope.) The caveat to this analysis is if, for example, this particular wave has a period of five minutes (as might result in the higher frequency AR in Figure 38), that corresponds to a phase speed of 0.9 m/s and a single phase will cross the entire field of view in approximately 45 seconds. Because the effective exposure time is 54 seconds, the effect of the phase lag would be mostly washed out. On the other hand, a single phase of the 40-minute oscillations seen in Figure 37 would take nearly six minutes to cross the FOV if due to the above wave, corresponding to more than a six frame, or 3.6 minute phase lag (which is not evident in Figure 37). Cross-correlations of residuals from CCDs at opposite ends of the SDSS FOV both along rows and along columns also indicated no phase lag in either RA or Dec residuals (Figure 39). The time difference between when a star is observed by CCD r'_1 and CCD g'_1 (288 seconds) is corrected during processing by offsetting the frame numbers in each of the second through fifth rows of CCDs so that the first frame in each run, for example, corresponds to the same time on all CCDs, rather than the same RA. This correction is not included in Figure 35 (hence the apparent lag between CCDs), but is made for all subsequent figures through Figure 39.

The unique operation mode of SDSS with data taken from a wide range of declinations, allows us to study the relationship between AR and airmass. If AR is caused by atmospheric waves at some range of altitudes above the telescope (*e.g.* 500-1500 km) we hypothesize that the amplitude of the AR residuals will increase with zenith angle and hence with airmass ($airmass \cong \sec z$, where z is the zenith angle). The increase would be due to an increase in the angle at which the waves are viewed, as well as a possible

91

increase in wave depth through which we were observing. Figure 16 indicates that we should see a much larger refraction effect when viewing at 45 degrees as compared to observing at the zenith (Figure 14), even when only a single layer monochromatic AGW is the source. Observations made by FASTT (Figure 4) also support this hypothesis, as they observed the amplitudes of the residuals observed at a zenith angle of 70 degrees to be roughly twice that observed at 45 degrees.
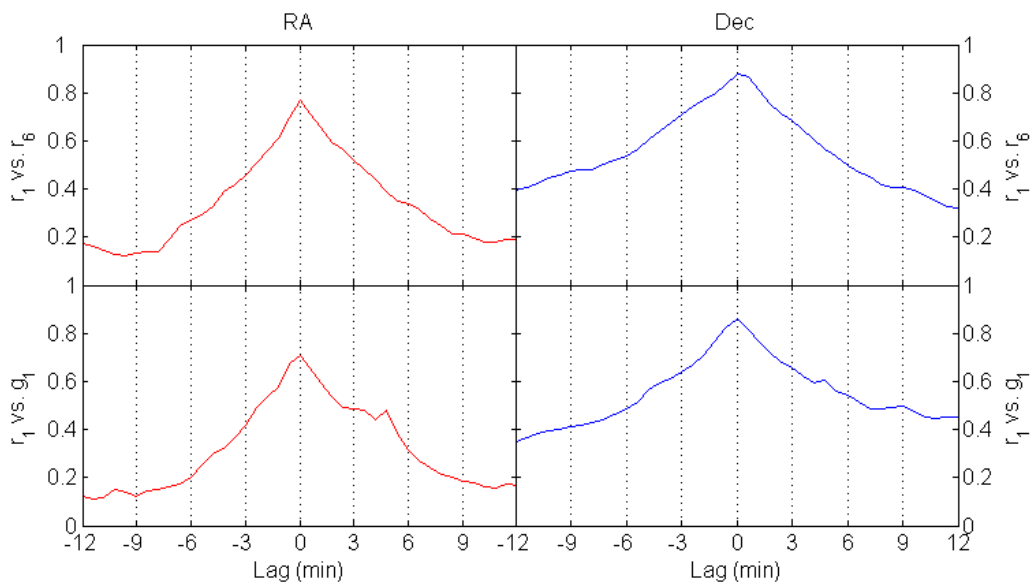
Figure 39. Normalized cross-correlation analysis between CCD's at opposite ends of the field of view (stripe 82 27-Sept-05). Top: opposite *r'* CCDs (across FOV perpendicular to scan direction). Bottom: first *r'* and *g'* CCDs (across FOV parallel to scan direction).

We compared SDSS residuals from runs taken at high, middle and low zenith angles, to look for any dependence on airmass (Figure 40). The top three plots in Figure 40 show residuals vs. frame number for each run and smoothing spline fits to the low frequency motions. The bottom three plots are the residuals from each run with the plotted fits subtracted to show the high frequency motions. Neither the high, nor low frequency residuals indicate any clear relationship with the airmass. Because all of the data was

taken at zenith angles smaller than about 53°, it's still possible that there is a relationship between AR and zenith angle, as observed by FASTT for very large angles. However, FASTT observed AR that nearly doubled in amplitude when observing at a zenith angle of 70° verses a zenith angle of 45° (Figure 16) so we might expect to see a fairly significant change in amplitude across the range of zenith angles.
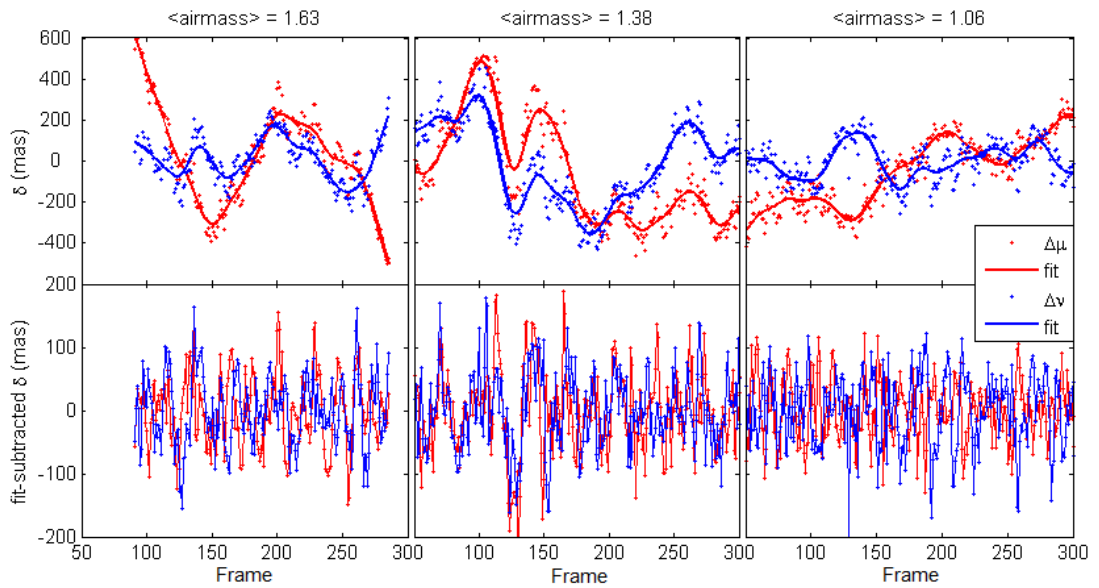
Figure 40. Comparison of residual obtained under different airmass conditions. Top: mu and nu residuals and smoothing-spline fits. Bottom: smoothing-spline subtracted mu and nu residuals.

The other implication of this analysis is that telescope tracking does not appear to influence anomalous refraction. Because the rotation velocity of the telescope axes increases with the angle of inclination of the great circle stripe relative to the equator, we would have expected residuals from stripe 44 (85° inclination), for example, to be larger than those from stripes 10 or 82 (0° inclination) if tracking had been a source of the error. In Figure 40, the first frame on the left is from stripe 44 (85° inclination), the middle frame is from stripe 82 (0° inclination) and the right frame is from stripe 24 (35° inclination).

The final comparison using the SDSS data was to compare residual characteristics to the weather conditions under which they were obtained. We were particularly interested in whether certain wind directions (*e.g.* winds incident from the direction of a nearby mountain) resulted in larger residual amplitudes. The primary relationships examined were how the residuals varied with surface wind direction and the effect of wind speed on residual characteristics.

We took the residuals and wind data from each of the 25 selected SDSS runs and binned the data into eight minute (two degrees of longitude or 13 frames) segments with each segment overlapping the next by half the bin width. For each bin the mean wind speed and direction is determined. The standard deviation of the AR residuals in each bin quantifies the amplitude as a function of time of those residuals varying on few minutes timescales. Binned wind conditions at the Apache Point Observatory during the selected runs are shown in Figure 41a, while the relationship between residual amplitude and wind direction is shown in Figure 41b. There is no obvious relationship between the amplitudes of the short period residuals and the wind direction. Comparison with binned wind speeds (Figure 42) clearly indicates that wind speed is not a factor in the short period motions.

Because there are significant components to the residuals with longer than 8 minute periods, we also performed the same analysis as above, but using 32 minute bins. The results, as illustrated in Figure 43, are characteristically the same as for the smaller bins.

94
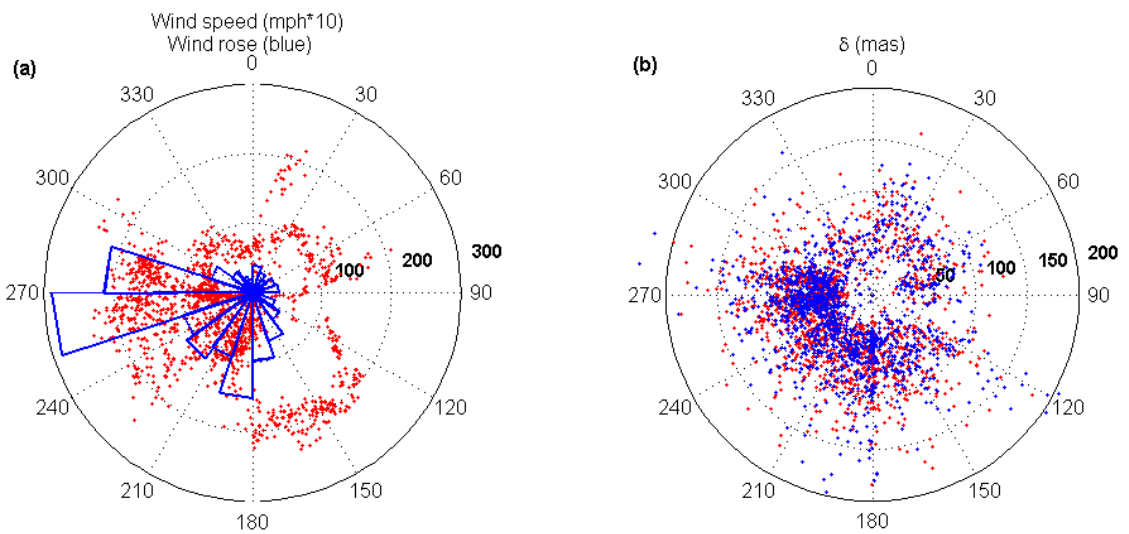
Figure 41. Wind dependency of SDSS data. (a) Wind speed*10 vs. direction (red points), and wind direction counts (blue). (b) µ (red) and ν (blue) residuals vs. wind direction. Wind data are means per eight minutes. Residual data are standard deviations of residual motions per eight minute bin.
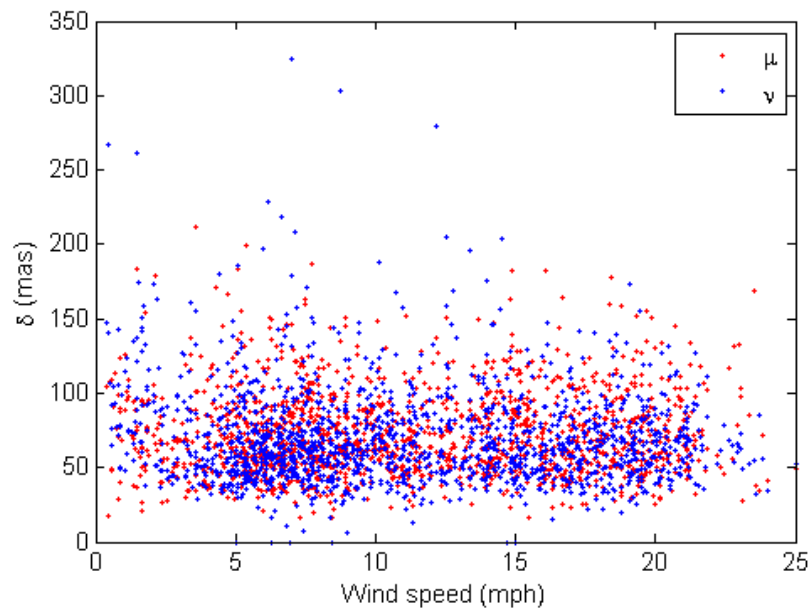
Figure 42. Standard deviation of residuals per eight minute bin vs. wind speed.
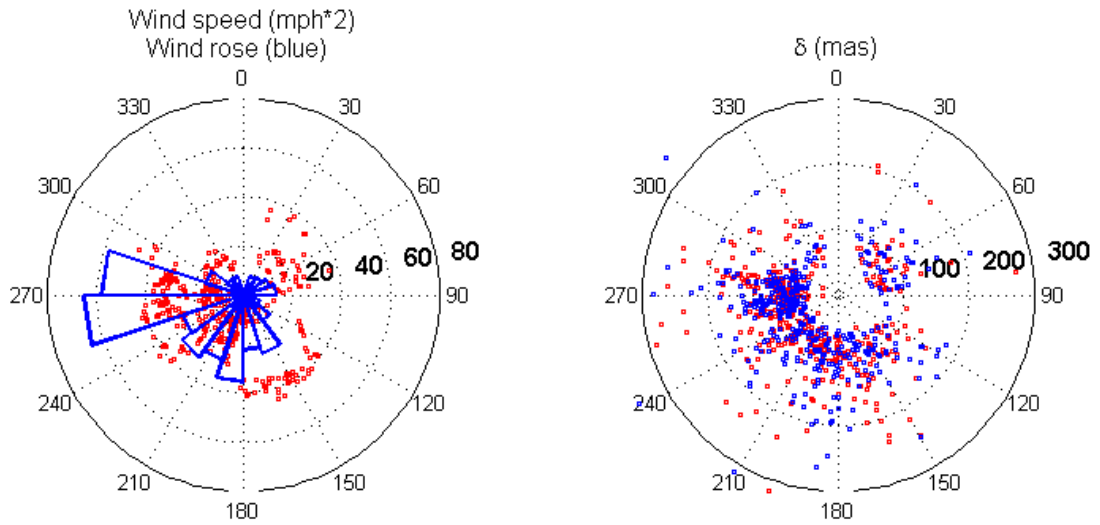
Figure 43. Same as **Figure 41** but with 32 minute bins.

To summarize the SDSS analysis, we re-reduced 25 runs, from both equatorial and high inclination stripes, examining residuals from all CCDs, as well as telescope metadata and wind conditions. Residuals consistently show the motions indicative of AR with periods ranging from minutes to hours and amplitudes of tenths of arcseconds to about one arcsecond. Row and column residuals are highly correlated across the focal plane, although widely separated CCDs show slight differences when comparing residuals. Despite the modulation of the residuals across the focal plane, there is no indication of phase lag, suggesting that while the source of the observed AR may change across the field of view, it is either not moving with any appreciable speed or is modulating much faster than it takes to cross the FOV. Comparisons between runs on different great circle stripes indicated no dependence on airmass/zenith angle or telescope tracking. There also was no clear color dispersion when comparing residuals from different filters and parts of the focal plane. Finally, both the short and longer timescale residuals have amplitudes that are independent of wind conditions.

96

3.3 AERI

The analysis of existing atmospheric data centers on the archives of the Atmospherically Emitted Radiance Interferometer (AERI) instrument available from the Atmospheric Radiation Measurement (ARM) program of the U.S. Department of Energy (www.arm.gov). AERI is a ground-based upward looking interferometer that measures the spectral intensity of atmospherically emitted radiation between 3 and 19 microns (infrared, IR) on eight minute timescales. The data products of primary interest to this research are the profiles of temperature and water vapor mixing ratio from the surface to 3000 meters which are calculated based on the radiance measurements.

As its name suggests, AERI is a Fourier transform spectrometer operating by means of a four-port Michelson interferometer, with one port directed at the zenith sky and two ports connected to reference blackbodies (Knuteson *et al.* 2004). The fourth port is covered so as to emit constant low levels of thermal IR radiation which are removed in processing. The blackbodies provide known temperature thermal IR references, one at ambient temperature and one at 330 K, both monitored using NIST (National Institute of Standards and Technology)-traceable thermistors. The use of two references allows determination of the slope and offset of the linear instrument response as a function of wavelength (Demirgian & Dedecker 2005). During an observation, a total of 200 seconds are spent observing the sky, with the remainder of the eight minutes spent observing the blackbodies. The blackbody calibration scans are made every two minutes allowing absolute calibration accuracy to better than 1% (Feltz *et al.* 2003).

Standard Michelson interferometer operation involves a movable mirror which alters the path length difference between the observed light source and a reference, resulting in variation in the intensity of the combined beam as a function of phase difference. For a non-monochromatic source, the superposed intensity is a sum of the intensities of the combined beams at each wavelength and each component intensity varies at a different rate as a function of wavelength when the mirror is moved. Based on knowledge of the mirror position as a function of time, the output temporal signal can be decomposed into its component wavelengths and intensities. AERI acquires an uncalibrated atmospheric scan (mirror sweep) every 2 seconds (Feltz *et al.* 2003) and the calibrated spectral radiance value output every 8 minutes is the average of approximately 90 interferometer scans (Knuteson *et al.* 2004).

An example of the AERI spectrum is shown in Figure 44. The measured intensities of the bands indicated in the figure provide the end points for the atmospheric radiative transfer equation. Using Radiosonde atmospheric profiles at the AERI site (acquired as often as every few hours) as a first guess, an iterative scheme is applied to determine the temperature and water vapor profiles which best produce the observed spectral radiances (for more details see Smith *et al.* 1999). The derived profiles are height resolved to 100 meters in the first kilometer of the atmosphere, 200 meters between one and two kilometers and 250 meters from two to three kilometers. The AERI radiance does not contain information about temperature and water vapor profiles above three kilometers (Feltz *et al.* 2003).
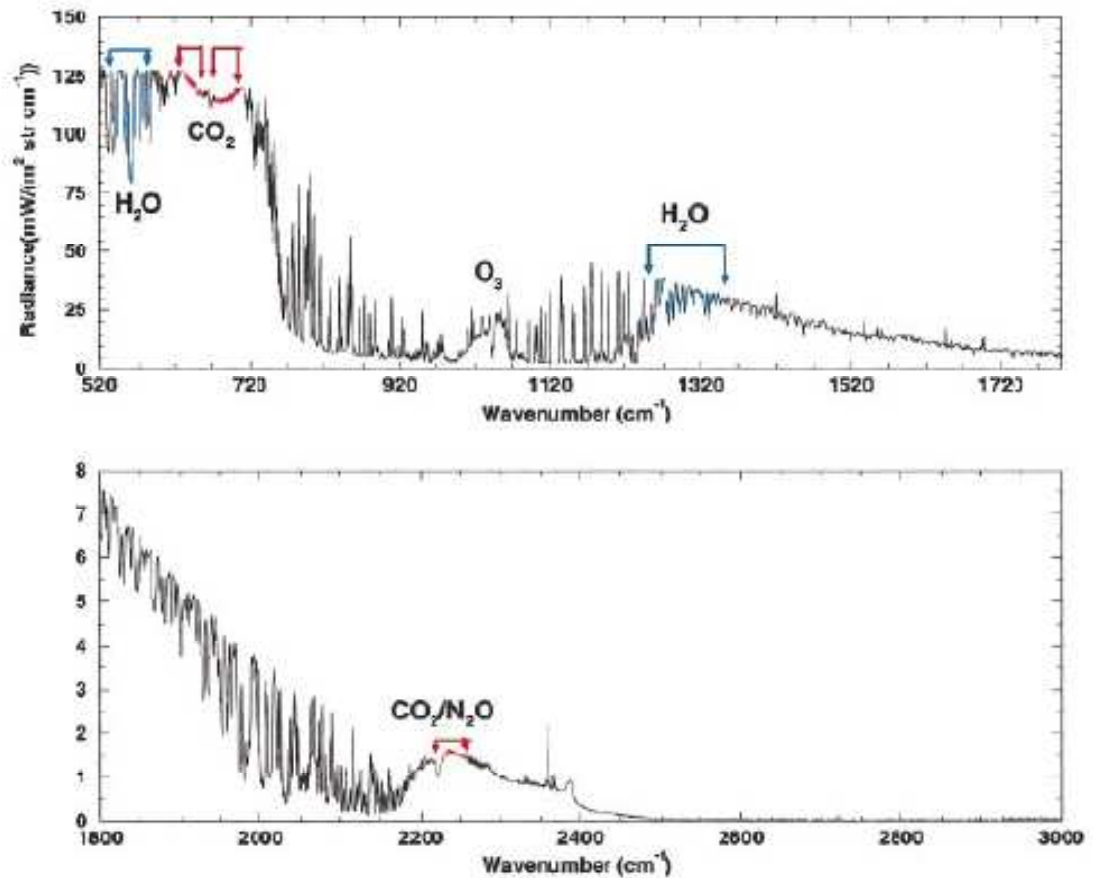
Figure 44. AERI radiance spectrum with regions used for temperature (red) and water vapor (blue) profile retrievals indicated. Spectra are plotted as a function of wavenumber in inverse centimeters (from Feltz *et al.* 2003).

3.3.1 Data Analysis

AERI instruments are (or have been) operated at several sites throughout the US and abroad, with the most suitable location (based on proximity and general atmospheric temperature and humidity conditions) for this research the Southern Great Plains site in eastern Oklahoma. While an AERI is no longer operated at this location, data were taken continuously for several years up to 2004. We chose approximately a dozen days of AERI data evenly distributed throughout the year in which most of the nighttime data was not degraded by cloud cover. (AERI operates under both clear and cloudy conditions

only ceasing operation with precipitation.) For each day, we limited the data to nighttime readings and flagged any datapoints taken under cloud cover. Using the AERI temperature, pressure and water vapor mixing ratio profiles and Equations (19) through (22) we calculated a grid of the atmospheric refractivity as a function of time and altitude.

An example of AERI profiles and calculated refractivity is shown in Figure 45. The temperature profile clearly shows the formation and strengthening of a boundary layer inversion beginning after 20:00 local time and extending from the surface up to 200 meters. The fractional change in refractivity is based on a standard atmosphere model anchored to the mean surface temperature, pressure and water vapor values for the night in question. As discussed in Chapter 2, standard tropospheric temperature decreases linearly with altitude, while pressure has an exponential dependence. Because the standard atmosphere assumes dry air, we used the average water vapor mixing ratio as a function of altitude over the selected nights of AERI data as the "standard" value. The downfall of using this standard atmosphere model is that the nighttime boundary layer temperature profile does not generally decrease linearly with altitude as is evident in Figure 45. A moderate improvement in the model was achieved by replacing the standard temperature gradient with a mean value obtained from all of the selected datasets. Because the mean temperature gradient contains only good nighttime data, it nicely represents the typical boundary-layer inversion. The model subtraction is used only for the purposes of visualization, and primary interest lies in the variability of the refractivity with time (the model is constant), so the issue of suitability of the model was not further addressed.

100
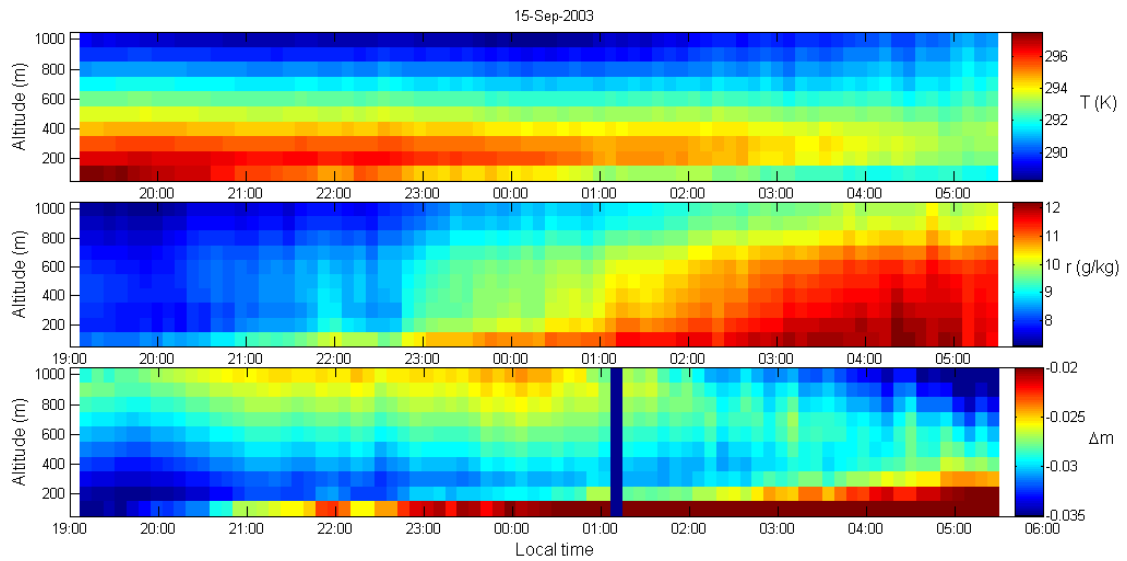
Figure 45. Example of AERI derived temperature (top) and water vapor mixing ratio (middle) profiles and calculated fractional departure of atmospheric refractivity from a standard atmosphere model (bottom).

Figure 46 shows the fractional departure from the model of refractivity for several nights throughout the year. The typical refractivity varies by as much as a few tenths of a percent on ten minutes timescales at a given altitude with changes as large as a few percent over the course of the night. The greatest diurnal variability occurs during the winter and summer. The dark blue vertical lines are times where data were discarded, either due to cloud cover or instrumental issues. For a typical boundary-layer refractivity value of 0.00023, a tenth of a percent would correspond to a change in refractivity of 2.3×10^{-7}, which based on Figure 13 would not likely be observable. The superposition of refractions from each layer of the atmosphere; however, may add up to a refraction we can observe.
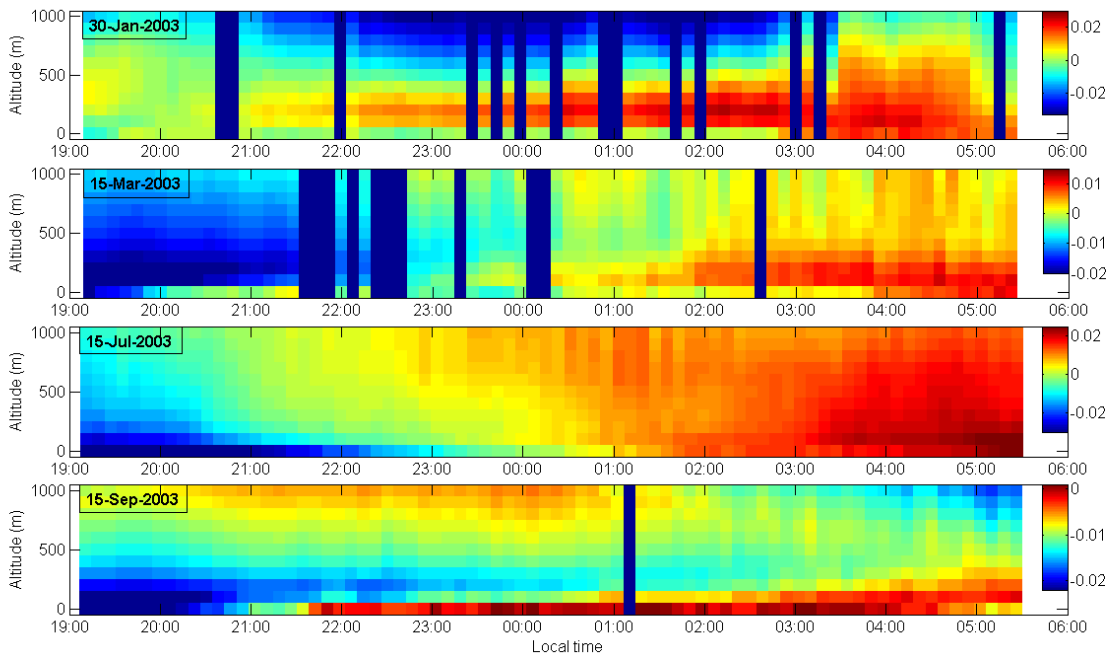
Figure 46. Four examples of AERI derived fractional departure of refractivity from a model. Color scale represents $\left(m(i,j) - m_{model}(j)\right)/m_{model}(j)$.

Using the AERI derived refractivity and the anomalous refraction model of Equation (24) we can calculate the theoretical refraction due to the boundary layer conditions observed by AERI. For this model we need to determine the slope of the refractivity gradient (*i.e.* the tilt of the atmosphere) at each point on the grid. Because AERI is a point source instrument and does not measure horizontal spatial scales of the atmospheric fluctuations it observes, we must define a suitable conversion from time to horizontal position. If we make the crude assumption that Taylor's (1938) hypothesis applies, *i.e.* that fluctuations are "frozen" in the atmosphere over the length of time it takes for them to pass over a sensor and travel with the mean wind, then we can use the wind speed to determine a horizontal scale. The applicability of this approach is limited because Taylor's hypothesis generally only applies to small scale atmospheric turbulence, not the mesoscale structures observed in the refractivity data, which are most likely not moving

102

with the local wind. If we assume that the observed atmospheric structures are never moving *faster* than the local wind speed, then the use of Taylor's hypothesis provides an upper limit to the horizontal scales of the structures and thus a lower limit to the resulting refraction (a smaller horizontal scale means a greater tilt).

AERI does not provide wind speed measurements, but concurrently with AERI operations, the National Oceanic and Atmospheric Administration (NOAA) operated a wind profiler demonstration network (WPDN) at the same location. The WPDN consists of 404 MHz clear-air radars operating at a wavelength of 74 centimeters (Martner *et al.* 1993). The radars detect turbulent fluctuations in the index of refraction from the surface to the tropopause (about 16 km) and obtain horizontal wind speeds by tracking the motions of turbulent structures with time. The profilers produce wind speed measurements from 500 m altitude up to 16,250 m with a 250 m altitude resolution and a six minute time resolution. Surface winds are obtained using an anemometer on a (~five meter) tower next to the radar.

For each night we find the magnitude of the wind velocity at each altitude as a function of time. The speeds are averaged over the night to find the typical speed for each altitude and then interpolated to a grid with 100 meter altitude resolution to match the boundary-layer AERI data. For each time bin of the AERI data, the horizontal spatial scale at altitude is defined as the temporal separation between the bin in question and the adjoining bin multiplied by the wind speed at that altitude. The vertical scale of the atmospheric tilt at a point on the time/altitude grid is the horizontal change in refractivity

103

between the point in question and the next time bin divided by the vertical refractivity gradient at the point, *i.e.* $\Delta y = \Delta m_h/(dm/dy)$. The resulting tangent of the atmospheric tilt is $\Delta y/\Delta x = \Delta m_h/[\Delta x(dm/dy)]$.

3.3.2 AERI Results

Examples of the results of this analysis are shown in Figure 47 and Figure 48. In both of these examples the amplitude of refraction that results from the conditions observed by AERI is at least an order of magnitude smaller than what we see with CTI and SDSS. The eight minute time resolution of AERI makes it impossible to observe periods of less than 16 minutes, which unfortunately is one of the chief domains of anomalous refraction. We do see periods of a few tens of minutes as in astrometric observations. The two primary explanations for the small amplitudes are errors in the model (*e.g.* the conversion from times to horizontal spatial scales as discussed above or incorrect reasoning regarding refractive index structures and atmospheric tilt) or simply that the atmosphere as observed by AERI is not the source of anomalous refraction.
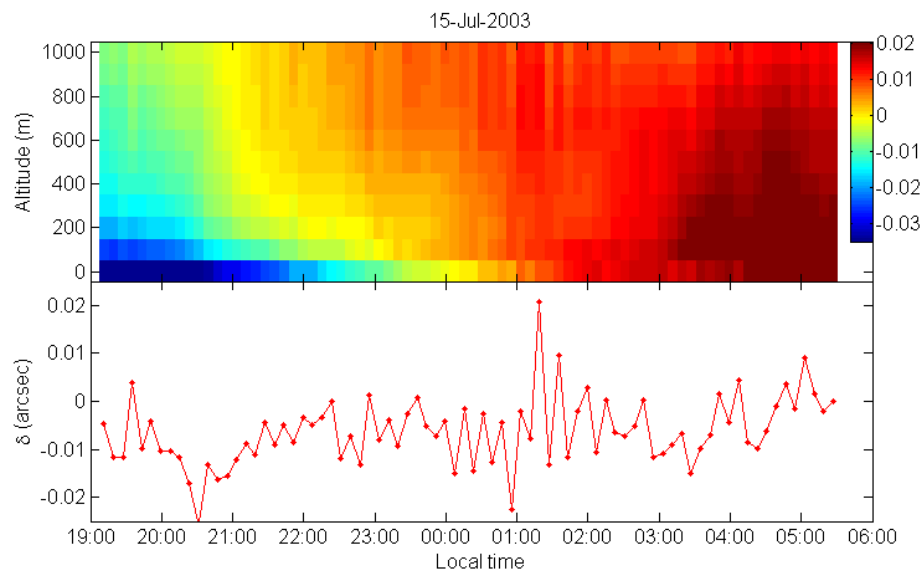
Figure 47. Fractional change in refractivity from model and calculated resulting anomalous refraction.

104
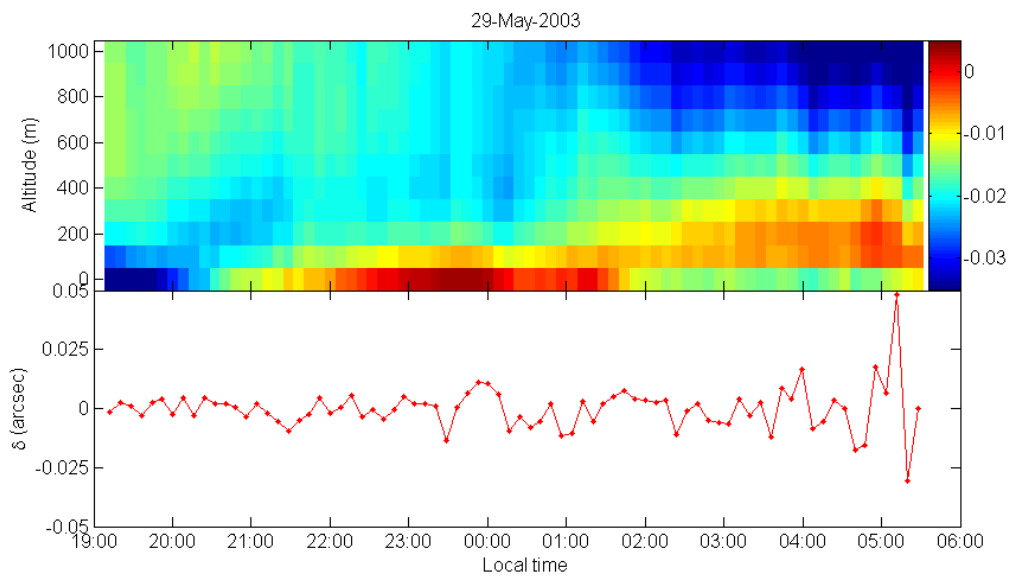
Figure 48. Same as **Figure 47** for a different date.

Another consideration when considering AERI profiles and the derived refractions is the spatial resolution of the AERI results. In the lowest kilometer of the boundary-layer AERI provides a resolution of 100 meters vertically. Atmospheric gravity waves and other potential sources of anomalous refraction likely have amplitudes ranging from a few meters to at most a few hundred meters. An isolated single layer disturbance with amplitude much less than 100 meters has been found theoretically capable of producing refractions of up to several tenths of an arcsecond (see the left side of Figure 21 or Figure 126 in Chapter 6). If such a source is the primary factor in AR, the resolution of AERI precludes observation of the resulting refractions.

105

4. Observations

Observations for this research were designed to determine the origin of anomalous refraction (AR): whether AR is instrument independent, to ascertain the angular and spatial scales over which AR is coherent and to clarify the atmospheric conditions under which AR occurs. These science drivers set the requirements that observations be made over several nights with variable atmospheric conditions on a minimum of two, preferably non-identical telescopes separated by no more than a few hundred meters. For the base goal of disentangling AR from instrumental motions, the telescopes must be operated simultaneously in drift-scan mode, pointed so that their fields of view (FOV) overlap. For the secondary goal of comparing the occurrence of AR with atmospheric conditions, said conditions, including surface winds, temperatures and sky conditions should be monitored continuously throughout observations.

Availability of instrumentation allowed observations to be made using three telescopes separated by 50 to 300 meters for multiple nights as well as two telescopes separated by two meters for many additional nights. In tandem with the telescopes, atmospheric observations were made using a specially designed microbarograph array which provided information about low altitude pressure variations and surface wind and weather conditions were continuously recorded with a weather station.

The need for multiple telescopes is driven by the hypothesis that any telescopes observing the same part of the sky from the same location should observe the same anomalous refraction regardless of differences in telescope design. This is not only a test

of the atmospheric nature of anomalous refraction, but will also aid us in separating telescope effects from anomalous refraction effects.

4.1 Astrometry

Observations were made over the course of three years on a variety of telescopes yielding a total of more than two dozen nights of astrometric data, and nearly as many with atmospheric data. The observations can be divided into three primary groups, a three night engineering run at the US Naval Observatory, Flagstaff Station in Arizona (NOFS) in 2007, 9 nights of science observations at NOFS in 2008 and an additional 15 nights of observations at the UNM Campus Observatory in Albuquerque (UNMCO) in 2008 and 2009.

The goals of the 2007 run were to determine and address the logistics of parallel telescope TDI operations with the NOFS telescopes because this was not a previously used technique at the observatory, as well as development of the data reduction routines required to see anomalous refraction in the data. The 2008 NOFS observations were centered on determining if AR is coherent on the scale of the observatory as well as if there is any relationship between weather conditions and AR. Simultaneous microbarograph observations were intended to show any correlation between AR and AGWs. The 2008-2009 UNMCO observations were devised after the results from the NOFS observations ruled out the original hypothesis (see Chapter 5). These observations were intended to further constrain the spatial scales of AR as well as to elucidate its source. The microbarographs were not operated during the UNMCO operations as will be discussed in Chapter 5.

4.1.1 Observatories and instrumentation

As the primary center for astrometric observing in the United States, the NOFS was an ideal location for the first half of this research. The observatory is located in the mountains to the west of Flagstaff in an area topographically suited for the production of atmospheric gravity waves (adjacent mountains being considered prime generators of boundary-layer AGWs as described in the Chapter 2). The FASTT telescope mentioned in the introduction is located at NOFS and is one of the modern TDI telescopes that have observed anomalous refraction. This makes the possibility of observing anomalous refraction at the observatory likely. There are currently four telescopes in operation at the observatory: in addition to the FASTT, which is an 8" meridian transit scanning refractor, there is a 1.55-m folded Newtonian, and two Ritchey-Chretien telescopes with diameters of 1.3-m and 1.0-m. All of these telescopes can be operated in drift-scan mode in the zenith as required by our observing program. They are all located on the same mountain within a distance of less than half a kilometer, which should allow observation of the same wave structures with minimal phase offsets under the assumption (Stone *et al.*1996, Pier *et al.* 2003) that AR is caused by AGWs with kilometer scale wavelengths. The spatial separations of these telescopes are great enough to sample a single AGW wavelength and thus allow definition of the spatial coherence of the anomalous refraction phenomenon.

The Naval Observatory (see Figure 49) is located on top of a low mountain at an elevation of approximately 2250 meters, characterized by two rounded peaks a few hundred meters apart separated by a shallow saddle. The vegetation on the mountain is

primarily coniferous forest with a canopy top at a few tens of meters and minimal undergrowth. The surrounding topography consists of rolling mountainous terrain of generally lower altitudes than the observatory. The San Francisco Peaks are an isolated group of high mountains located approximately eight miles to the northeast of the observatory with a maximum elevation of 3800 meters. The 1.3-m and the 1.55-m telescopes are located on the western peak of the NOFS mountain at a separation of approximately 50 meters, while the 1.0-m telescope is located approximately 250 meters to the ENE of the 1.55-m on the eastern peak. The 1.0-m telescope has a slightly higher elevation than the two larger telescopes, but all telescopes are located on similar terrain.

Figure 49. Aerial view of the USNO Flagstaff showing the three telescopes used and their relative positions.

The 1.0-m telescope is housed in a large dome which is connected to a single story building (housing the control room and a machine shop) by an open-air walkway covered

by an old rectangular, peaked roll-off observatory roof. The roof is slightly higher than the top of the telescope. The only other structures in the vicinity of the 1.0-m telescope that are of the same height or taller than the telescope (and thus a potential source of orographic effects) are trees surrounding the observatory building at no closer than a few tens of meters. To the northeast of the 1.3-m telescope are the main office building of the USNO Flagstaff and the dome of the 1.55-m telescope. The 1.55-m dome is considerably taller than the 1.3-m telescope and may be a significant source of surface layer disturbance. The 1.3-m is also in the vicinity of tall trees to the south and west. The 1.55-m telescope is housed in a large dome on the second story of the main USNO Flagstaff office building. As such, the telescope is considerably higher than all nearby structures, although the hill to the northeast (on which the 1.0 m is located) may be a source of "orographic" activity.

While the two smaller telescopes are Ritchey-Chretien optical systems, they otherwise have little in common. The 1.0-m telescope was built in 1934 by George W. Ritchey and has a 7.3-m focal length. The 1.3-m telescope was built by DFM Engineering and was designed for wide field (degrees) observations, having a 5.2-m focal length. The detector on the 1.0-m telescope is an e2v CCD with 2048X4100 13.5 micron pixels, providing a field scale of 0.38 arcseconds per pixel and a CCD field of view of 13x26 arcminutes. The 1.3-m detector package has a mosaic of six e2v CCDs, each chip having 2048X4102 15 micron pixels, providing a field scale of 0.595 arcseconds per pixel and a CCD field of view of 20x41 arcminutes per CCD. The 1.3-m focal plane array is shown in Figure 50. The serial register is along the short dimension of the CCD on all of the NOFS detectors.

110
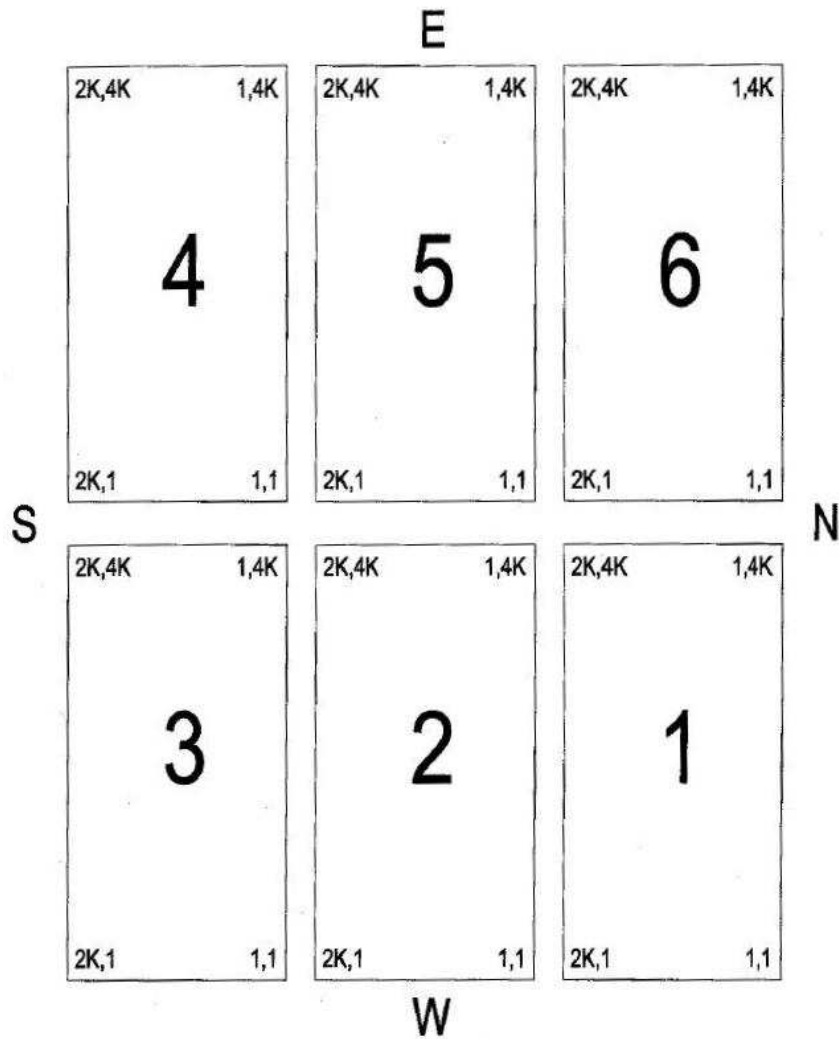
Figure 50. Focal plane array of the 1.3 meter telescope at the US Naval Observatory, Flagstaff Station. Perspective is looking up at the sky.

The 1.55-m Kaj Strand astrometric reflector is a folded Newtonian with an effective focal length of 15.2 meters. The telescope was designed for accurate astrometry on small fields and has a field of view of 11 arcminutes when operated with a SITE CCD with 2048X2048 24 micron pixels. The field scale is 0.325 arcseconds per pixel.

The University of New Mexico Campus Observatory (see Figure 51), site of the observations made in late 2008 and early 2009, is located on the north end of the UNM

111

campus on a small hill at about 1520 meters elevation. The Sandia and Manzano Mountains, with elevations in places above 3000 meters, are a more or less continuous ridge, at the nearest approximately nine miles east of the observatory and run north-south for several tens of miles. The building is centered in a large paved parking lot with the landscape within a few meters to roughly ten meters on all sides sparsely populated with vegetation typical of the high desert environment (small shrubbery, cacti, *etc.*). We used two 10 inch Meade LX200 telescopes for the observations, both of which are located at the west end of the observatory courtyard, several meters south of the main observatory dome. The main dome is on top of the observatory building and rises several meters above the domes of the two Meades. To the east of the telescopes are two additional domes (housing a spectrophotometer and a lidar), both of which are also slightly higher. The two telescopes are aligned approximately east-west and are separated by 2.25 meters.

Figure 51. Aerial view of the UNM Campus Observatory with locations of the two telescopes used in this research indicated, as well as the locations of the other domes at the observatory.

112

The two Meades are effectively identical and were built in 2005 with f/10 Schmidt-Cassegrain optical designs and 2.5 meter focal lengths. Each has a Finger Lakes Instruments ML6303 camera incorporating a Kodak KAF-6303E CCD with 2048X3072 9 micron pixels, providing a calculated field scale of 0.73 arcseconds per pixel. The true field scale is approximately 0.63 arcseconds per pixel with the discrepancy resulting from the focal length being longer than the published number. The serial register is along the long dimension of the chip.

4.1.2 Operations

For observations on the 1.3-m telescope we chose to use only one of the CCDs in the camera array. Obtaining science grade drift-scan images on all six CCDs was not possible in the chosen operational mode. Due to the alignment and clocking requirements of TDI imaging described in Chapter 2, we were only able to achieve suitably round point-spread-functions on one CCD at a time. In order to have all six CCDs producing science-grade images we would have had to align the camera rotator to a very high precision and clock each column of CCDs at a slightly different rate (which was not actually possible with this camera). The centers of each column of CCDs are separated north-south by nearly half a degree, with the result that a clocking rate that is correct for the central column of CCDs will be off by several hundred microseconds for the north and south columns. The images on those CCDs will be horizontally elongated by as much as ten arcseconds or more.

The CCD mosaic is aligned on as close to a grid centered on the meridian as feasible, which is good for stare-mode exposures. No CCD is on the meridian, however, which

means that for such a wide array the curvature of the motion of stars across the CCD array for zenith scans results in significant smearing of the images. By rotating the camera such that a single CCD is aligned with its center on the meridian and setting the tracking rate to the declination at its center, we can minimize the TDI artifacts on that CCD, but all other CCDs will have very large artifacts because they will not be aligned with their respective meridians and the tracking rate will be wrong for their declinations. Because the main science goal involved comparing residuals obtained simultaneously on the different telescopes, and the FOV of a single CCD on 1.3-m is greater than that of either the 1.0-m or the 1.55-m, this comparison only required science grade images on a single 1.3-m CCD.

The smaller fields of view and shorter exposure times (exposure time of the zenith pointing TDI images is set by the drift time of stars across the FOV) of the 1.0-m and the 1.55-m CCDs makes acquiring round TDI images on these two telescopes a considerably easier task. By far the roundest images were acquired on the 1.55 meter where the very short (54 second) exposure time coupled with the very small (11') field of view (and consequently small range in declinations) result in minimal TDI artifacts. The downside of the small FOV and short exposure times on the 1.55 meter is a reduced stellar density compared to the other two telescopes. The 1.0 meter with its 2048 column by 4100 row chip has an exposure time that falls between those of the 1.3 meter and the 1.55 meter resulting in slightly more noticeable TDI artifacts (some stellar elongation towards the top and bottom edges of an image), but still rounder images than the 1.3-m.

The one technical issue with the set up of the 1.0-m and 1.55-m cameras was the restriction to clocking at 100 µsec resolution. A 100 µsec difference in clocking rate at a declination of approximately 35 degrees corresponds to a difference in declination where that clock rate matches the sidereal rate of more than 16 arcminutes. In other words, if a clocking rate of 31100 µsec per pixel matches the sidereal rate at a declination of 34°42', then a clocking rate of 31000 µsec will match the sidereal rate at a declination of 34°58'. During the second 2008 observing run we chose to slightly offset the declination of all of the telescopes from zenith to a declination corresponding to a 100 µsec interval tracking rate on the 1.0 meter and 1.55 meter. We then adjusted the clocking rate on the 1.3 meter (adjustable at the microsecond level) to match the new declination.

The two UNMCO telescopes were set up specifically for this research with small fields of view and comparatively short exposure times minimizing TDI artifacts. The primary complication in setting up these telescopes is the lack of a camera rotator. In order to align the cameras such that stars trail precisely down the pixel columns, the telescopes themselves must be rotated in azimuth. If the telescopes are pointed exactly at the zenith and the telescopes mounts perfectly level, a rotation in azimuth will not change the telescope pointing; however, any deviations from this perfect alignment will cause the telescope pointing to trace out a cone on the sky with changes in azimuth. For single telescope operations, the absolute pointing is not a major concern, but for parallel telescope observing, we need all telescopes to have both the proper camera rotation and to point at the same field on the sky (so that we can be sure they are looking through the same column of atmosphere). Prior to beginning parallel telescope operations we leveled

115

the telescope mounts as accurately as possible and determined that the telescopes could be pointed at the zenith with an accuracy of a few arcminutes, allowing at least a partial overlap of the fields of view regardless of azimuthal rotation.

All data were taken in continuous drift-scan mode and broken into contiguous frames corresponding to the size of the CCDs used. Thus the 1.3 meter frames are 2048x4102 pixels, the 1.0 meter frames are 2048x4100 pixels and the 1.55 meter are 2048x2048 pixels. The exposure times were 198 seconds on the 1.3-m frames, 124 seconds on the 1.0-m and 54 seconds on the 1.55-m. For both UNMCO telescopes the frames are 2048X3072 pixels with exposure times of 105 seconds.

We began the first set of telescope operations the night of the 10th of June 2007 (UT date) at NOFS and were able to almost immediately begin taking data on the 1.0-m telescope (drift-scan operations, including rotation and tracking rate had been tested prior to our arrival and the logistics worked out). The only operational problems with the 1.0-m were the use of an incorrect filter (b' instead of r') during the first part of the night and the occasional tendency of the camera computer to crash, necessitating a reboot. Data were taken ten frames at a time during the nights of June 10 and 11 as that appeared to prevent the computer from crashing. No data were taken on the night of June 12 due to weather. The June 13 data were taken in sets of 100 frames, except less when a computer crash occurred.

Drift-scan operations on the 1.3-m telescope had not been tested prior to the observing run and proved more difficult than expected. The first two nights of 2007 observations were spent determining the correct combination of camera rotation and tracking rate for science-grade images on any one of the six CCDs. On the night of 12-13 June we achieved the best images of the run on CCD 3 (see Figure 50) and took science data for the whole night. Despite turning the telescope drives off once the telescope was pointed, the telescope pointing still had a tendency to drift over the course of the night.

Unfortunately the data obtained on the 1.3-m telescope suffered from severe TDI artifacts despite all attempts to obtain the best clocking rate and camera alignment. Most notably, stars off the optical axis of the frame were consistently smeared into "kidney beans" (see Figure 52), making the actual positions of the center of each star very difficult to pinpoint. The resulting errors in astrometry far exceeded the tolerances for observing anomalous refraction. An additional problem with the 1.3-m telescope frames discovered during processing is that there is no overlap with the fields observed by the 1.0-m telescope. During the observing run we failed to consider that the telescope pointing coordinates correspond to the center of the CCD mosaic, not the particular chip which we were using. Because we used chip 3 (see Figure 50), the actual coordinates to which we were pointing were over 20 minutes of declination below where the telescope drive encoder stated we were pointing, resulting in no declination overlap between the frames from the two telescopes. In the end, after several failed attempts at extracting useful information from the 1.3-m data, primarily due to the image distortions, the data were

117

deemed unusable and discarded. While the lack of overlap in the telescope pointing wouldn't have rendered the observations useless, poor image quality did.
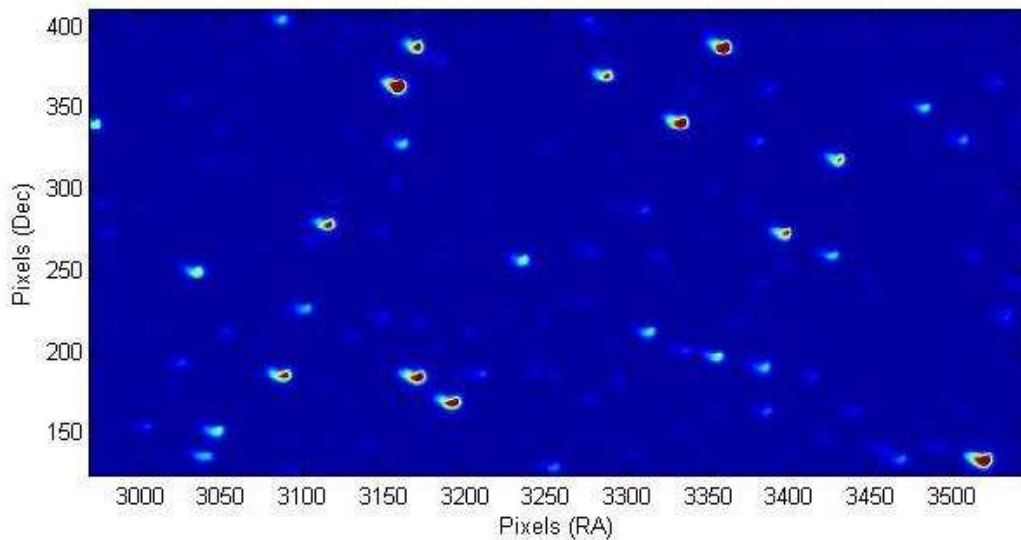
Figure 52. Subsection of a frame from the 1.3-m Ritchey-Chretien taken during the June 2007 observing run at USNO Flagstaff showing the distortion due to TDI smearing.

In 2008 we were granted ten nights of time on both the NOFS 1.0-m and 1.3-m telescopes as well as four nights of time on the 1.55-m. These nights were divided into five nights from April 26 – April 30 on the two Ritchey-Chretien telescopes with two nights on the Strand reflector from April 29 – April 30 and another run of five nights from May 10 – May 14 on both Ritchey-Chretiens with two nights on the Strand from May 13 – May 14. Of the observing time we were granted we lost only one full night and a few hours from one other night to weather and one full night on the 1.3-m and a few hours on all the telescopes to mechanical issues.

The first 2008 observing run started out much like the 2007 observing run. The 1.0-m had a new software system that alleviated the crash problem of the previous run and

118

started up with no obvious initial problems. Acquiring the proper alignment and clocking rate for the 1.3 meter took most of one night of observations but we ultimately were able to achieve far better images than in 2007. Both telescopes were determined to be centered on the same field of view.

Unfortunately, part way through the first night of observation we discovered a new problem with the 1.0-m with the potential for rendering much of the data taken on the telescope useless. Examination of frames showed that some of the images exhibited what appeared to timing glitches. In the more severe cases, a glitched frame would have double images of most or all of the stars (see Figure 53). The problem was traced to the CCD control software (newly implemented just prior to this observing run) which was being interrupted in the process of clocking the CCD. The interruptions resulted in the clocking of the CCD faltering such that charge collecting from a given star would be late in advancing to the next pixel and would therefore lag behind the image of the star. If the delay was small this might result in an elongated stellar image or in the more severe cases double images of stars. If the clocking was interrupted many times the result was a "chain of pearls" effect with multiple offset images of the same star. Initially this effect was only apparent in the most severely affected frames where the delays were largest; however, closer examination of the timestamps on each row indicated that these glitches were occurring on average over a hundred times per frame, with most of the delays being of order a few milliseconds (see Figure 54). The normal tracking rate for the 1.0-m at zenith is 30800 μsec per row, so the average frame would have glitches corresponding to a total delay over the course of the frame of order 20 rows.
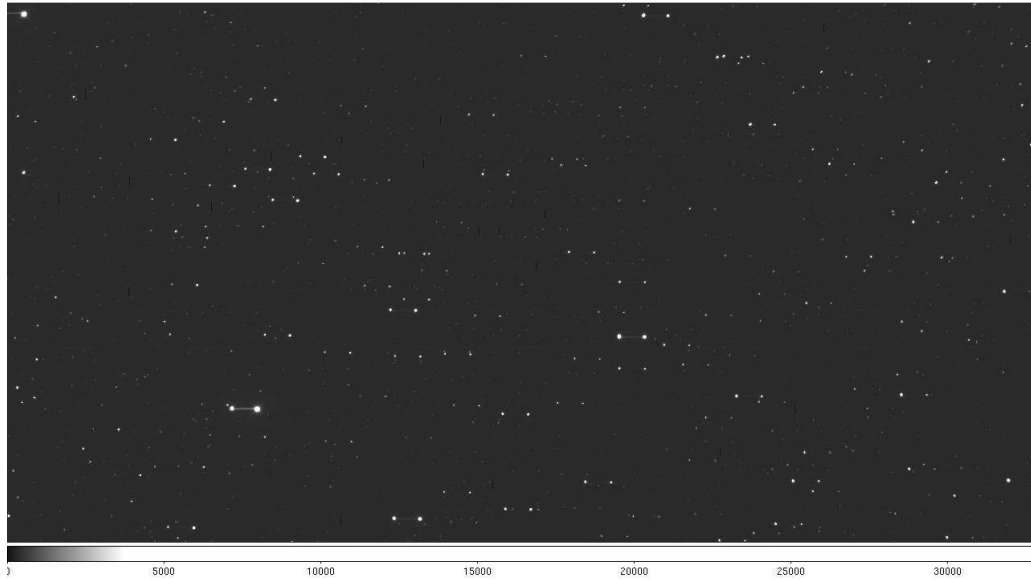
Figure 53. April 2008 1.0-m frame exhibiting a severe TDI timing glitch resulting in double images of the stars.
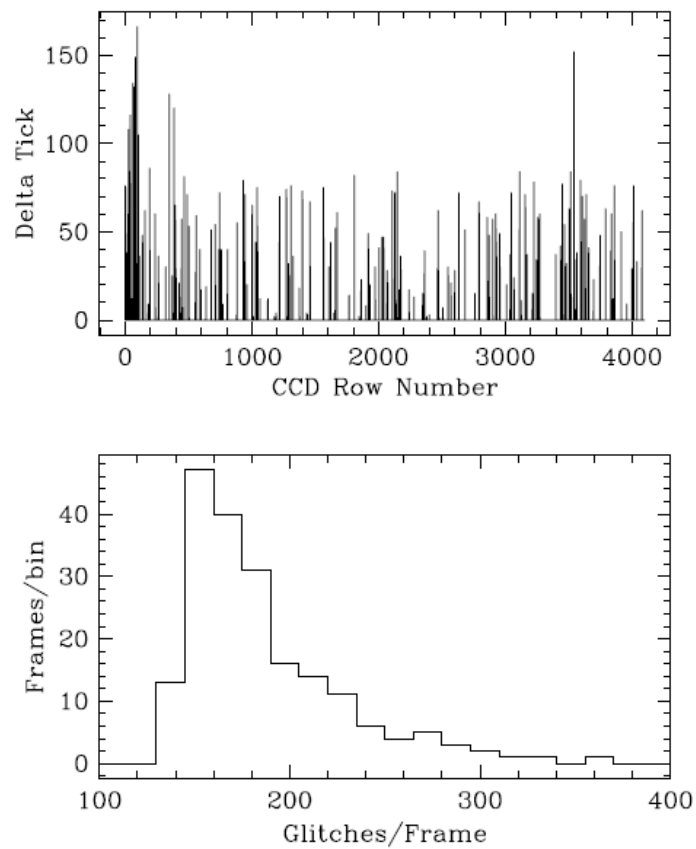
Figure 54. 1.0-m April glitch statistics. Each tick is 0.0001 seconds so a timing glitch of 100 ticks corresponds to an error of 10 milliseconds or 0.3 pixels (0.114 arcsec).

120

In the latter part of the April run we attempted to minimize the glitch count by shutting down as many other software processes on the CCD control computer as feasible, but did not find this to have a significant impact on the number of glitches per frame. A number of other solutions to the glitch problem were proposed, none of which could be implemented during the April observing run.

The final result of the April 2008 observing run on the 1.3 meter was four nights with approximately four to six hours of science data each and one night with a little more than an hour of science data. All of the 1.3-m data suffers from moderate TDI artifacts; however, the level of distortion is not so severe as to prevent centroiding with useful precision. From the 1.0 meter we obtained five nearly complete nights of science data, but the presence of timing glitches has compromised all of these data to varying extents. Finally, we obtained two nearly full nights of data from the 1.55 meter, again most of which is contaminated by timing glitches. Because no suitable solution was found to reliably remove the glitches and because the timing glitches not only hamper our ability to match stars to a catalog, but also may mimic AR in RA, these data were unsatisfactory for this research.

The May 2008 observing run was by far the most successful of all three NOFS runs. With our recent experience on all of the telescopes we were able to quickly get both the 1.0 meter and 1.3 meter telescopes up and running early on the first night with minimal adjustments required. In between the April and May observing runs a new observing protocol on the 1.0 meter had been implemented which nearly eliminated the timing

glitch problem that plagued the April run. The combination of a faster CPU on the CCD computer and running the CCD control software under root privileges with high priority cut the number of interruptions in the CCD timing to just a few per night. A few frames still experienced glitches of a hundred or more ticks and were removed from the science data. The remaining frames had no more than a few glitches of a few ticks each amounting to an error of less than a hundredth of a pixel (a discussion of frame clocking and ticks will follow in the next section), or less than a milliarcsecond – well below the centroiding precision we can achieve. The instrument rotator alignment on the 1.3 meter was slightly better on this run than the previous as well. All of the telescopes ran smoothly for the entire observing run and weather allowed us four full nights of the five that we were allotted, one of which included operations on the 1.55 meter.

UNMCO operations began in November of 2008 with several nights of data taken between the end of November and mid-December on a single 10" Meade telescope, 10"E (the second Meade, 10"W, not yet operational). These initial operations were approached with the goal of determining whether these telescopes were stable enough and could see enough stars with high enough signal to noise ratio to extract useful information regarding anomalous refraction. A lidar (see Section 4.2.3) was operated simultaneously during several nights of astrometric observations allowing comparison between residuals and the low atmosphere lidar returns.

Operations resumed in April of 2009 with the commissioning of 10"W. The parallel telescopes were operated continuously, as weather (which was uncooperative) and

instrumental issues (one night lost to a failed power supply) allowed between April 30 and June 30. Ten nights with anywhere from two hours to 7.5 hours of data were obtained with the two telescopes pointed at the same field on the sky. One night was obtained with 10"W tilted by $0°.68$ towards the east and two nights were observed with 10"W pointed $1°.16$ towards the east. Care was taken to ensure that the declinations of the two telescopes were the same during these latter operations so that the precise relative positions in RA could be determined based on the time lag between when a particular star was observed by 10"W and when it passed through the field of 10"E.

In order to tilt the 10"W telescope towards the 10"E telescope (so that the fields of view of the two telescopes cross at a lower altitude in the atmosphere allowing potential pinpointing of the altitude of the source of AR) it was necessary to tilt the base plate to which the telescope was mounted. The Altitude-Azimuth mounting of the telescopes requires a $90°$ rotation of the telescope (so that it's aligned E-W) in order to use the drives to tilt the instrument towards the east. Tilting the mounting plate instead allows the telescope to maintain its north-south orientation required for the stars to track precisely down columns on the CCD.

Five additional nights were obtained with the telescopes angled away from each other to test the outer limits of the angular scales of AR. Two nights were obtained with the east telescope pointed half a degree north or south of the 10"W declination, one night was obtained with 10"E pointed $1°.5$ north of 10"W and for two nights 10"W was pointed ten

degrees south of 10"E. These pointings were made using the telescopes' altitude drives with the mounting plates horizontal.

4.1.3 Astrometric data processing

The purpose of the astrometric data reduction is to take many hundreds of astronomical images and reduce them to a considerably smaller volume of data directly applicable to this experiment, hence the term reduction. In particular the reduction aims pinpoint the positions of the stars in each frame, find a suitable plate transformation to match the stars to a standard reference catalog to determine celestial coordinates and positional offsets from the catalog (residuals); and extract the stellar statistics (magnitude, full width at half-maximum, shape, *etc.*), image quality metadata and instrumental metadata.

The reference catalog is approached from several angles – existing astrometric catalogs, catalogs formed from a compilation of the image data and direct night-to-night and/or telescope-to-telescope comparisons. Each of these approaches has certain advantages and disadvantages when applied to this project.

For the astrometric catalog comparisons we primarily use the UCAC2 catalog (Zacharias *et al.* 2000) and the Carlsberg Meridian Catalog, release 14 (CMC14, Copenhagen Univ. Obs. *et al.* 2006). We also considered using the Naval Observatory Merged Astrometric Dataset (NOMAD, Zacharias *et al.* 2004), a compilation of astrometric data from several catalogs, including UCAC2, with the most precise data for each star (if data from more than one catalog is available) listed. NOMAD has the advantage of a much higher density of stars than UCAC2, but the catalog data for many of the stars has large

astrometric errors (several tens to hundreds of milliarcseconds). After comparing NOMAD residuals to UCAC2 residuals, we determined that the errors in NOMAD were dominating any anomalous refraction we might see. Creating our own reference catalog from our observations allows us full knowledge of errors and conditions of the observations. The downfall is the limited number of observations of each star (as well as the very narrow window in which those observations were made) and the smaller time span over which we have observations from all nights. Finally, there is the basic approach of directly comparing observations from one night to the next and from one telescope to another. This approach does not give us absolute information on anomalous refraction, but does tell us if time-dependent refraction effects were observed on different nights or if different telescopes saw the same refraction effects on the same night, all of which is valuable information for this research.

4.1.3a NOFS Data

The first step of processing the NOFS data (see Figure 55) is to determine an approximate central RA and Dec for each frame and adjust the headers and image names accordingly. Each frame is actually 2112 (for the 1.0 m and 1.55 m) or 2116 (for the 1.3 m) columns wide with the extra 64-68 columns divided between overscan and underscan regions and several columns containing metadata about the telescope during the scan. The overscan region indicates the bias level of the CCD as a function of time while the metadata columns include the time that each row is read out along with other CCD stats (CCD temp, *etc.*). Time is recorded in units of integer ticks (starting at zero at UT midnight) with the precision of the time record determined by the number of ticks per second, which is dependent on the computer operating system used.

125

Image processing pipeline
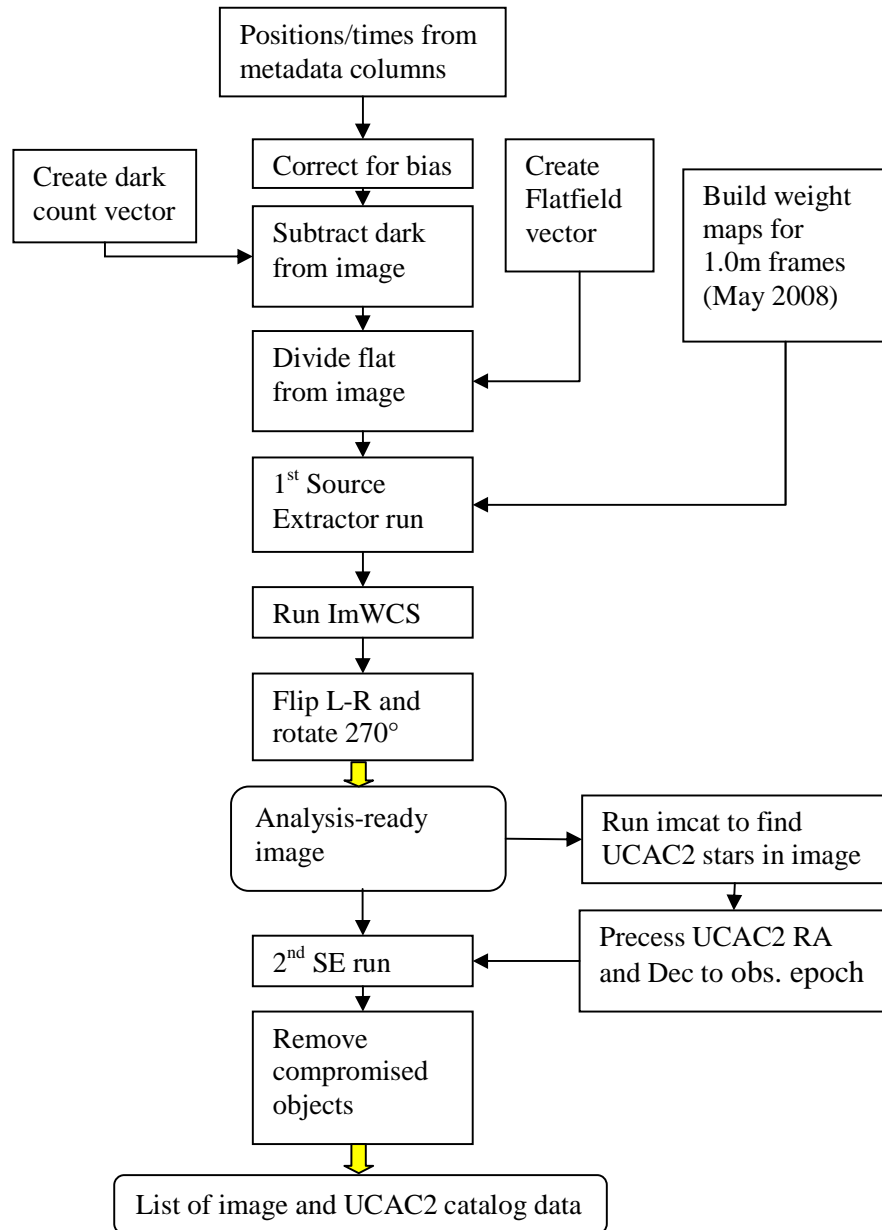
Figure 55. Image processing flowchart for 2008 observing run data. The primary end-stage products, indicated by the yellow arrows, are the analysis-ready images and the lists of image and catalog data.

For the 1.0 meter there are 10000 ticks per second, providing a 0.1 millisecond precision, while on the 1.3 meter there are 5000 ticks per second. For each row the tick count is recorded in two 15 bit columns. One column contains the number of multiples of 2^{16} while the other ranges from zero to 2^{16}-1. When the low-order column reaches 2^{16}, it flips back to zero and the high-order column increments by one. The tick count is calculated by multiplying the high-order column by 2^{16} and adding it to the low-order column. The time at which any given row was read out can then be calculated from the row's tick count (see Appendix B, code *getrowtime.m*).

Read-out times of the central row in each frame for each telescope are determined from the time metadata columns and converted into the local apparent sidereal time (LST), or RA. The one exception to this is the 1.0 m data taken during the 2007 observing run because timing data were not recorded in the metadata columns. For those frames the approximate frame time is taken from the observing logs. The resulting frames are renamed to include the central RA, date, telescope and filter for each image in the image name. This information, as well as the central UT time of the frame, epoch, plate scale and frame number are all appended to the fits header of each image (see Appendix C, code *prep13_1.m*).

At the start (and occasionally end) of each observing night dark current frames were taken by "exposing" an image for the length of a standard drift-scan frame with shutters and dome closed, such that the detector is unilluminated. Several standard zero-time exposure biases were also taken, but proved to be identical to the ramp-up scan mode

dark frames. For drift-scan observing the ramp dark frames and the zero exposure biases are not ideal. In the first frame of each set of exposures (the ramp frame) the exposure time of each row increases linearly with row number, causing the number of counts to also increase linearly with row number. This occurs because charge is only clocked across a fraction of the frame dependent on row number. In the science images the charge from each star is clocked across every row in the detector, thus if there is a hot pixel (a pixel with a higher than normal rate of charge leakage) in the same column as that star in any row of the detector, the extra charge produced by the unilluminated pixel will be added to the signal of the star. In a ramp or a zero bias frame the charge from a hot pixel is only clocked across part of the frame (i.e. from the pixel's position on the detector to the end of the frame), so while on a science frame or non-ramp dark a hot pixel will appear as a bright column, a ramp or zero bias will show a hot pixel as a some fraction of a hot column dependent on where the pixel is located in the detector. Subtracting this from a science frame with a full hot column would only partially correct the hot column. In short, the bias structure seen by a TDI image is the pixel bias structure averaged over each complete column (because a given star "sees" each row for an equal fraction of the total exposure time). For these reasons only the scan mode darks were used for correcting the images.

The overscan region of each frame (including the darks) provides an estimate of the DC offset of the bias for each row. In every frame (dark, flat and science) prior to any other correction we fit a line to the overscan counts as a function of row and then remove this trend from the frame itself, correcting any change in bias counts with time. Each night's

dark frames (typically ten non-ramp frames) are concatenated into a single 2048 x N superdark. A sigma-clipped mean is then found for each column to create a one-dimensional dark vector which can then be subtracted from each row of each image (Appendix D, code *bias.m*). The dark vector is the same average pixel bias that is seen by each row of a science image, with each point in the vector corresponding to a particular column in the images.

We derived a flatfield for each night directly from the science frames. Approximately five frames from each night taken after moon-set or before moon-rise (depending on the observing run to minimize gradients in the images) are bias and dark corrected and then concatenated into a single superflat frame as was done for the dark frames. A sigma-clipped histogram of each column is characterized by a normally distributed peak centered on the typical number of sky background counts with a rapid fall off towards zero pixels at higher counts (see Figure 56). By fitting a parabola to the logarithm of the sky background region of the histogram we can determine the mode of the distribution of sky background counts. By repeating this procedure for each column in the superflat we create a vector of the background as seen by a drift-scan image, *i.e.* the "average" background in each column over the course of an exposure (Appendix E, code *flat.m*). Any irregularities in the telescope or detector (dust, pixel irregularities, *etc.*) will change the overall background in a given column when seen in drift-scan mode. Dividing each row of the dark-subtracted science frames by the normalized flatfield vector ameliorates the signature caused by such telescope irregularities.
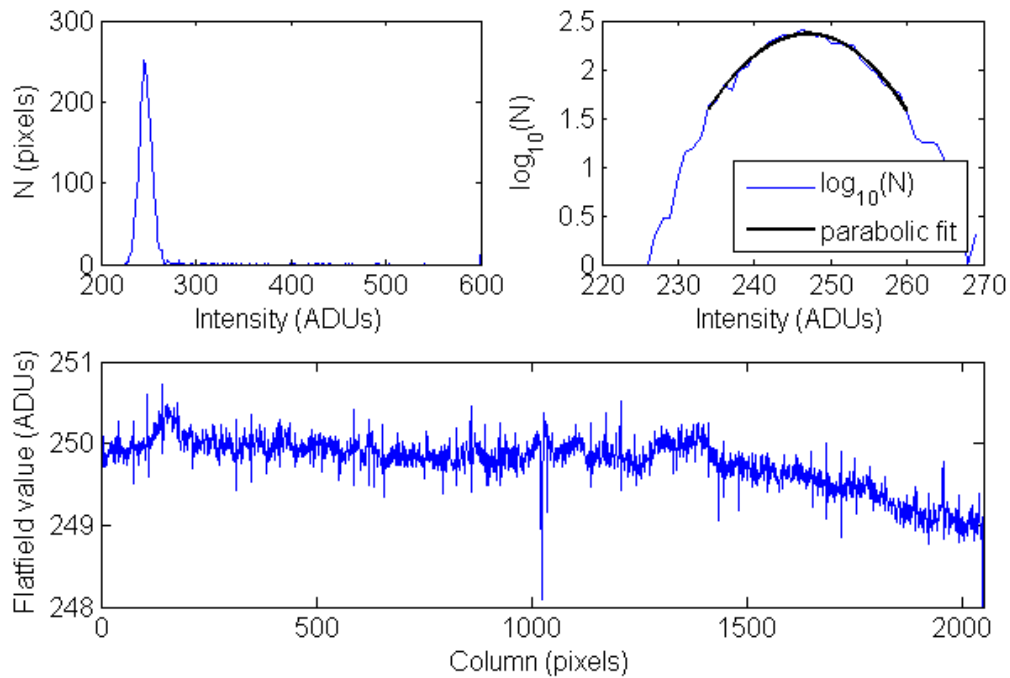
Figure 56. Flatfield creation: The upper left frame is a histogram of a single column from a 1.3 meter science frame (truncated to 200-600 ADUs). The upper right frame is a logarithmic plot of the peak of the intensity distribution with a parabolic fit. The intensity value at which the derivative of the parabola is zero corresponds to the modal night sky background. The final flatfield vector for this image is shown in the bottom plot.

Flats correct for differential sensitivity across the field viewed by the CCD. To assess that variation in sensitivity, we assert that the sky background evenly illuminates the FOV of the telescope. Under stare-mode operations flatfields are commonly created either by observing the (arguably flat) sky at dusk or by imaging an evenly illuminated screen. The supposition is that by evenly illuminating every pixel in the detector (through the same instrumentation and optical path as used by the science images) variability in the sensitivity (whether due to dust, pixel irregularities, *etc.*) will appear as variations in the image intensity. The flatfield for a TDI image must record the sensitivity variations as they are observed in TDI mode, at the same clocking rate as the science images. This sets the exposure time at a value suitable for nighttime

observations, but well beyond the saturation limit for dusk or illuminated screen images. The night sky background at a dark sky site under clear conditions is arguably flatter than the commonly used dusk sky. Because of the averaging nature of TDI mode requires a flatfield that is the average intensity of each column on an evenly illuminated detector, we can take the modal value of the counts in each column of a science image as the average background, thereby ignoring the presence of stars in the image.

A glitch in the image readout on the 1.0 meter telescope during the May 2008 observing run necessitates an additional image processing step. Due to what is believed to be another timing glitch with the CCD controller, occasionally pixels were read out as having zero counts. This typically occurred in groups of 32 pixels along a single row (zero count lines), with anywhere from a few to several hundred zero count lines in a single frame. On the large scale, these glitches should not affect the overall processing of the 1.0 meter data; however, in the isolated cases where a zero count line coincides with a star in a frame, the star's position and flux will be altered by the presence of the glitch. In order to prevent the data from being contaminated by the zero count lines, at the start of the image processing routine a map is created for each frame indicating the location of every zero count line in that frame (see Appendix F, code *seimw.m*). Later, this map is converted into a weight map that indicates to the centroiding program, Source Extractor (description follows), to ignore any star that coincides with a zero count line.

Once frames have been bias and dark subtracted and flatfield corrected they are written out to a new image which is reflected left-right, and then rotated 270 degrees to align the image coordinate axis with that of the sky as depicted by the astrometric catalogs used.

An initial list of pixel positions and magnitudes of stars and other objects in each frame is found using the centroiding program Source Extractor (SE, Bertin & Arnouts 1996). SE determines a mean sky background value for each image, locates and centroids the stellar positions. Elongated objects such as galaxies or convolved stars are flagged and ignored in later stages of processing. For the centroids we used SE's windowed positions, which are found by applying Gaussian window to a circular aperture for each object and finding the intensity-weighted mean position within that window. Photometric characteristics are calculated within an ellipse determined by an object's second order image moments. The current version of SE (2.4.4) has centroid variances that are typically of order $10^{-2} - 10^{-3}$ pixels^2, which is comparable to the centroiding precision of two other commonly used astrometry algorithms, DAOPHOT (Stetson 1987) and Photo (used by SDSS) (Becker *et al.* 2007). Additionally, comparing variances in the spline subtracted residuals (Figure 57) resulting from SE windowed positions fit to the Carlsberg Meridian Catalog (CMC14) to the same spline subtracted residuals using positions found by fitting a two-dimensional Gaussian to each stellar PSF shows that the PSF-fit positions are in general comparable to or worse than the SE positions, but produce the same AR signatures (Figure 58).
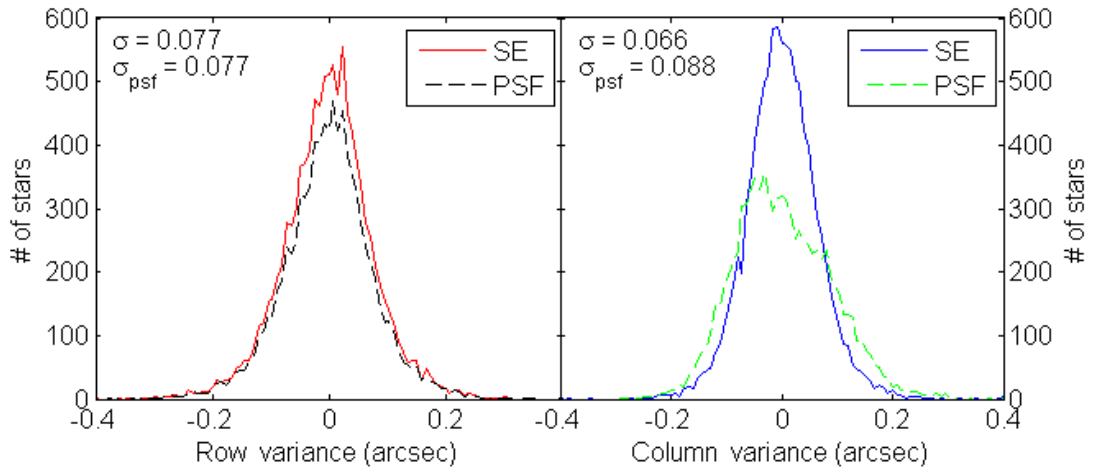
Figure 57. Comparison of Source Extractor windowed centroids to positions produced by fitting a two-dimensional Gaussian to each stellar PSF. Residuals are the difference between measured pixel positions (converted to arcseconds) and CMC14 catalog positions.
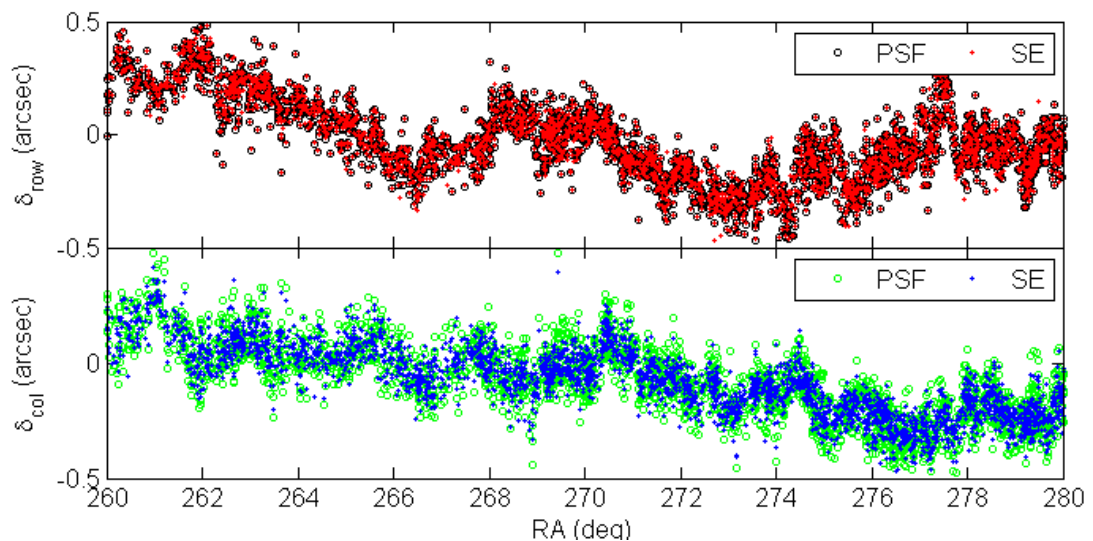
Figure 58. Comparison of residuals obtained using SE windowed centroids and those based on two-dimensional Gaussian fits to the stellar PSFs.

The resulting text table of pixel positions and associated fits files are input into the WCStools program *imWCS* (Mink 1997) (Appendix F, code seimw.m). *ImWCS* fits a quadratic polynomial to the images to find an initial transformation between the image pixel coordinates and UCAC2 world coordinates. The headers of each image are then modified to include standard World Coordinate System (WCS) keywords which can be

read by most FITS image viewers. This includes the RA and Dec of the optical axis of the image (as well as the associated pixel positions), the pixel scale (in arcseconds per pixel, with separate scales for x, y and cross terms) and the polynomial terms used in the transformation from pixels to RA and Dec.

At this stage in the image processing routine we have analysis-ready frames and no further changes are made to the images themselves. The remaining steps in the pipeline extract information required for the analysis from the images and format these data into lists and matrices that we can then analyze.

Running the WCStools program *imcat* (Mink 1997) returns a list of catalog stars (in this case UCAC2) located in the field of regard of each frame as well as their pixel positions based on the WCS transformation found using *imWCS* and magnitudes. Catalog RA and Dec are precessed to the observation epoch (*i.e.* 2008) using the US Naval Observatory's Vector Astrometry Subroutines (NOVAS, Kaplan & Bangert 2006). A second Source Extractor run locates image stars associated with UCAC2 stars in the *imcat* list and returns a list of windowed pixel positions, associated UCAC2 positions, flags indicating potentially compromised objects, as well as any additional desired image characteristics such as flux, background, image moments and stellar FWHMs. For each frame flagged objects, saturated stars, stars with low S/N, blended stars, extended objects and stars very close to the edges of each frame are discarded (Appendix G, code *prep13_2.m*).

To allow analysis of each night as a continuous list of stars (as opposed to fitting each frame individually to the reference catalog) we determined the precise time at which each star was read out by extracting the integer pixel read-out time from the metadata columns and adjusting that time based on the sub-pixel centroided position of the star. We then describe each star's image position using time and y-pixel allowing a simple, single conversion from time to RA and y-pixel to Dec for each night.

Analysis of the June 2007 data proceeded differently from the analysis of 2008 data described above due to the lack of timing data in the underscan columns. The best method of retrieving residuals was found to be matching stars in overlapping frames from the different nights of the 1.0 m data. In each frame the WCStools program *immatch* was used to locate USNOA-2.0 catalog stars in the image. The high density of stars in the USNOA-2.0 catalog allows almost every star in each frame to be matched with a catalog star. We next created a catalog of pixel positions, magnitudes, USNOA-2.0 IDs, and USNOA-2.0 positions for each night. The ratio of number of rows per degree of RA is known from the tracking rate of the CCD. By comparing pixel and (USNOA-2.0) RA separations between frame centers to the known ratio, any gaps between frames (i.e. the length of time and therefore number of rows) can be determined. The row pixel values of the stars in each frame are changed to reflect the frame's position in a continuous strip image for the whole night. The result is a catalog for each night that we observed, wherein pixel positions reflect the positions of stars on a strip that is 2048 pixels wide by N pixels long, where N is the number of rows in a continuous strip image that would be obtained by unbroken zenith scanning from the start of the night to the end of the night.

135

Matched stars in strips taken on different nights were found by matching the USNOA-2.0 catalog ID numbers. Residuals are then defined as the difference in pixel positions between the matched stars. On closer analysis, data taken on the night of June 9/10 were found to be consistently out of focus, thus only the nights of June 10/11 and June 12/13 were used in the final analysis. All 2007 residuals are derived from a direct star-star comparison of these two nights.

Residuals for the 2008 data (both runs) were determined by finding the transformation from pixel positions (time and y-pixel as described above) to RA and Dec for each star that resulted in the smallest spread in residuals when catalog positions (UCAC2 or CMC14) were subtracted from image positions. Transformations were manually determined for each night (using the expected plate scale or pixel rate as an initial guess) and were linear with cross-terms added as necessary.

The quality of the transformation is determined by fitting a smoothing spline to the residuals to remove systematics and creating a histogram of the remaining noise in the residuals (see Figure 59 and Figure 60) as well as calculating standard deviations. Typical values of the standard deviation are between 60 and 90 milliarcseconds. The noise is a combination of stochastic catalog errors (CMC14 has published positional accuracy of 35 to 100 mas, Copenhagen Univ. Obs. *et al.* 2006; while UCAC2 has a positional accuracy of 15 to 70 mas) and errors in the derived centroids. Given typical seeing blurred PSFs of 2-3 arcseconds we can expect the centroiding precision for a

signal-to-noise ratio of 100 to be of order a few tens of milliarcseconds. The transformations to catalog coordinates are therefore providing fits to UCAC2 and CMC14 that are well within the error bounds imposed by catalog and centroid precision.
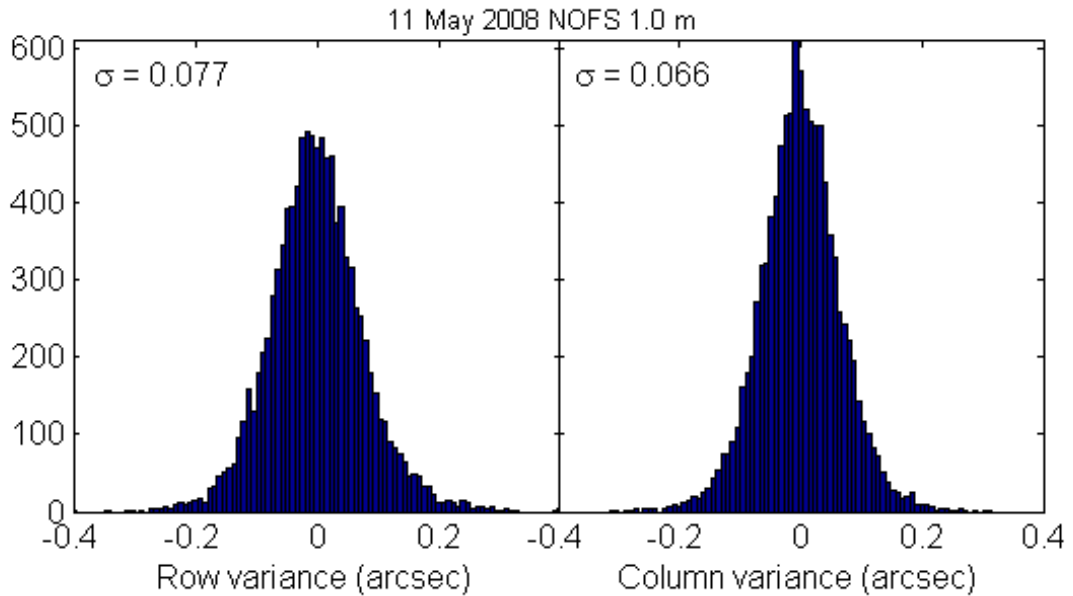
Figure 59. Histograms of spline-subtracted noise in row and column residuals (row – RA and column – Dec) for a linear transformation from pixels to RA and Dec for one night of May 2008 NOFS 1.0 meter data.
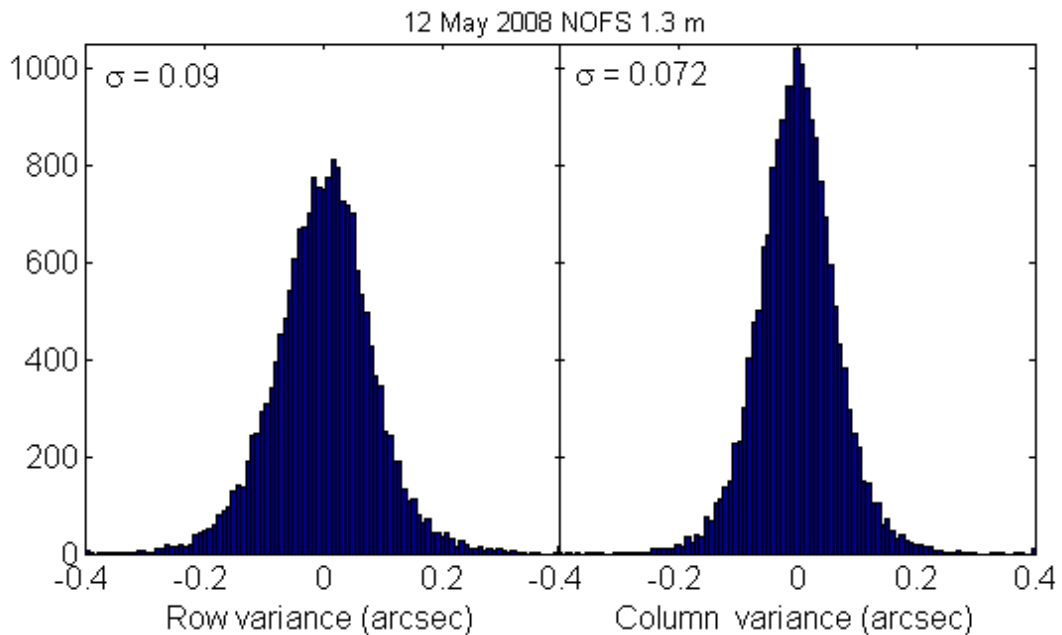
Figure 60. Same as **Figure 59** but for a night of 1.3 meter data.

4.1.3b UNMCO Data

Data reduction for the UNM Campus Observatory observations is very similar to that for the 2008 NOFS data. Raw images are named according to the UT date and time which they were read out and include encoder positions and times in the headers, greatly simplifying the initial data processing. The images do not contain overscan regions with bias levels so all bias removal is done with the dark subtraction. Starting with dark subtraction, the reduction process for UNMCO data follows the same procedures as with the NOFS data with a few minor differences. The images are rotated 90 degrees to align them with the catalog axes. Accurate timing information is not contained in the housekeeping columns, so as with the 2007 NOFS data, for row position we use pixel numbers adjusted to reflect an object's position as it would appear on a single continuous strip encompassing the entire night of observations. Stars are matched to both UCAC2 and CMC14 and row and column pixel positions transformed to RA and Dec using a linear fit that is found using the method of least squares, minimizing $RA - (c_1 + c_2 * X + c_3 * Y)$ and $Dec - (c_4 + c_5 * X + c_6 * Y)$, where c_i are constants, X and Y indicate row and column positions respectively and RA and Dec are the matched catalog positions. Generally a cubic polynomial is fit to the resulting positions as a function of RA to remove image motions with roughly diurnal timescales. These long-period motions are likely caused by temperature-induced structural changes, but may also contain a very low frequency element of anomalous refraction. In either case, these very long timescales are beyond the scope of this dissertation. Less often, an additional quadratic polynomial is subtracted from the Dec residuals as a function of Dec to correct for centroid biases caused by poor focus or camera alignment. Because these Dec errors are constant with

138

time, removing them will not affect the observed AR signatures. The quality of the fits is comparable to the NOFS data (standard deviations of 0.06 to 0.09 arcseconds, see Figure 61) despite the generally lower signal to noise ratios attainable with a small telescope in Albuquerque.
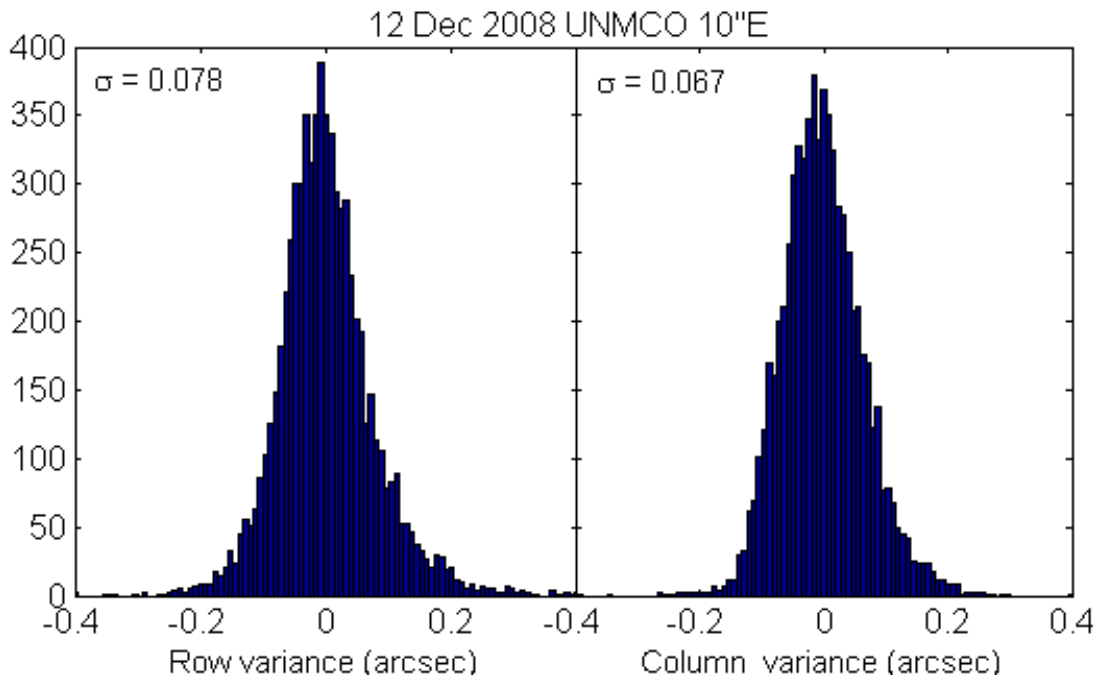
Figure 61. Same as **Figure 59** but for one night of data from the east 10-inch Meade at the UNM Campus Observatory.

4.1.4 Data Characteristics

On initial consideration of the astrometric data a number of unexpected characteristics were discovered. The first of these was the appearance of sub-integration time variations (an example is circled in green in Figure 62). The second, and more concerning, is the occasional significant increase in the spread of residuals at a particular RA ("blow-ups," see Figure 62, black circles).
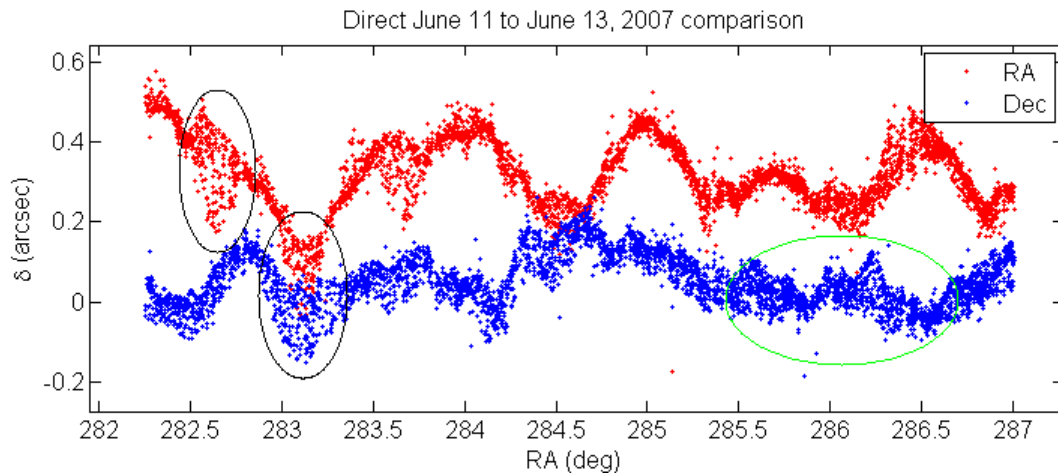
Figure 62. Contiguous overlap section of residuals from the 1.0 meter telescope. Systematic variations occurring on sub-integration timescales are circled in green and "blow-ups" are circled in black. A single integration time corresponds to approximately 0.5 degrees of RA.

It was originally hypothesized that any atmospheric fluctuations occurring on timescales shorter than the integration time (*i.e.* with periods of 124 seconds or less for the 1.0-m telescope) would be washed out. The drift-scan mode of operation is effectively a top-hat filter, with the observed position of each star being an average of the star's position over the 124 second integration time. If an atmospheric wave of period shorter than or equal to the integration time passes overhead during an exposure, a single star will be observed through one or more complete cycles of the wave. This was originally thought to result in a zero net perturbation; however, on further consideration it became clear that unless the period of the wave is an integer multiple of the integration time, the star will not see an integer number of cycles and the average perturbation will not be identically zero. This is precisely the result one expects to get from a top-hat filter convolution – periods shorter than the width of the filter are significantly damped, but not completely removed.

To clarify how this should affect the residuals we ran the basic experiment of simulating sine waves of frequencies ranging from much shorter than an integration time to frequencies of several minutes and convolving the waves with a top-hat filter with a width equal to the integration time. The result for two such sine waves is plotted in Figure 63. A key characteristic of the observed perturbation given a monochromatic sine wave input is that it is perfectly 180 degrees out of phase with the input waves for periods between half and one integration time and perfectly in phase for periods less than half an integration time. The observed perturbations for all waves of periods longer than the integration time will be fully in phase with the inputs.
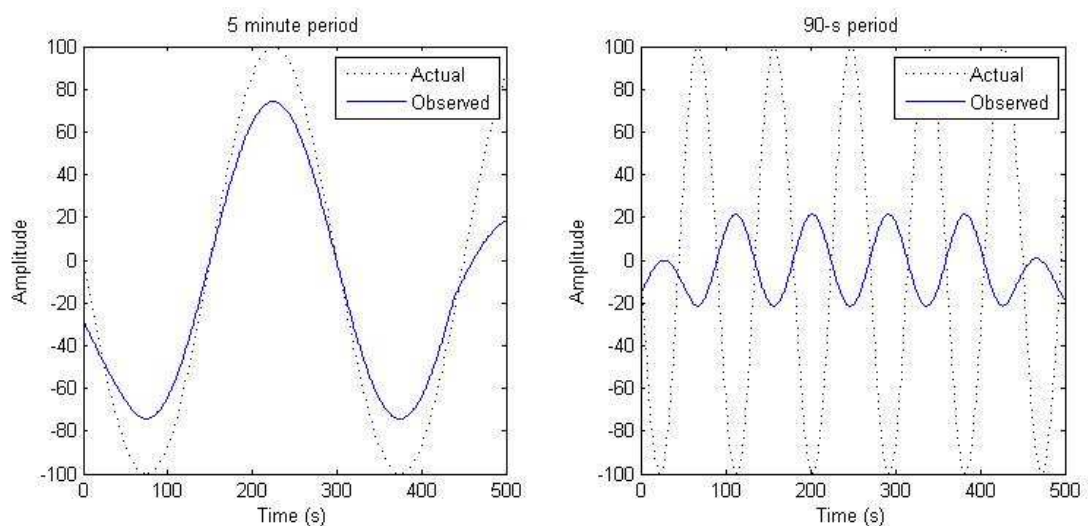
Figure 63. Simulated monochromatic waves and observed waves after convolution with a 124 second top-hat filter.

As an experiment more in line with what might actually occur in the presence of real AGW activity, we also examined the effect of convolving superimposed waves of different frequencies with the 124 second top-hat filter. The result is shown in Figure 64 and is a mess of perturbations in phase and out of phase with the input signal.

The key result from this particular analysis is that we have observed anomalous refraction occurring on timescales shorter than initially thought possible. The observed anomalous refraction is likely of much lower power than the superimposed source waves and will exhibit a varying phase relationship with the waves.
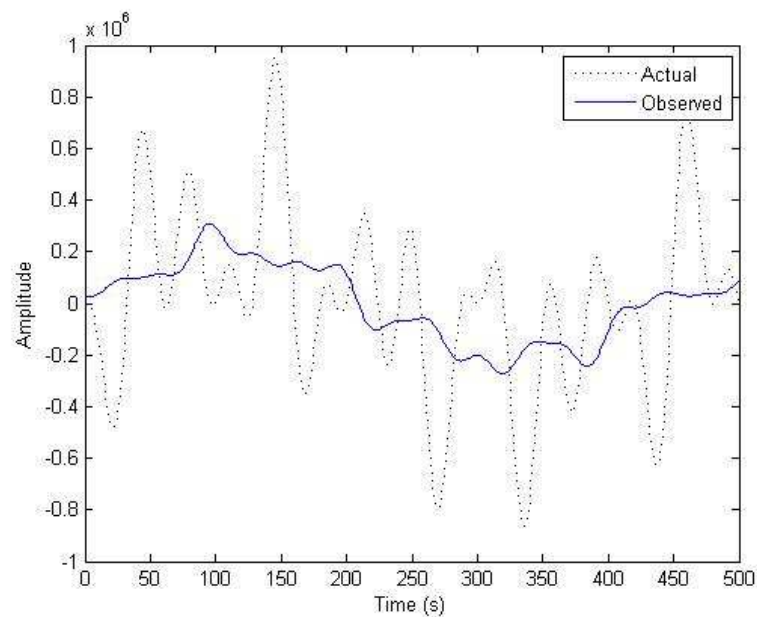
Figure 64. Simulation of observed perturbation (using a 124-s top-hat filter) given a superposition of waves with periods of 200 seconds, 120 seconds and 64 seconds.

Analysis of the blow-ups has proved to be much more complicated. We define the noise in the residuals to be any non-systematic positional differences, specifically, the width of the line of the residuals with the anomalous refraction removed. Blow-ups are regions of greatly increased noise and occur roughly every few frames and with duration of order the length of a frame (see Figure 65). The effect is seen most prominently in the residuals resulting from direct telescope to telescope (or night to night) comparisons and is seen consistently on all telescopes used (Figure 65 and Figure 66). Comparisons to

UCAC2 or CMC14 show little indication of the presence of blow-ups (Figure 67). In many instances the blow-ups are confined to the overlap region of two specific frames with the spread in the residuals quickly decreasing on either side of the overlap region. However, in other cases the blow-ups only occur in part of a frame overlap region. The occurrence of these blow-ups is more prevalent in the row (RA) than column (Dec) residuals and they tend to have larger amplitudes in RA as well.
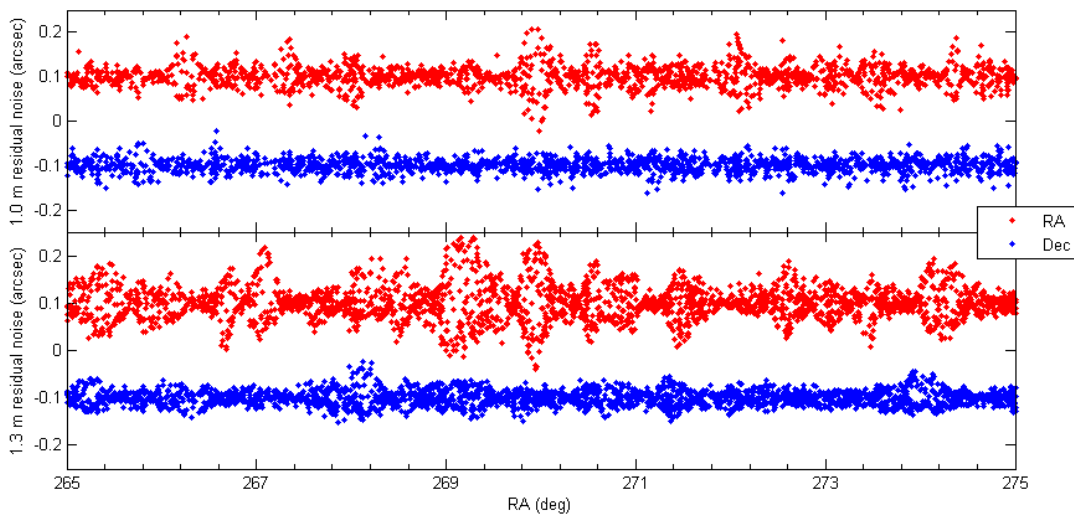
Figure 65. Spline-subtracted residual noise for both the 1.0 meter and 1.3 meter telescopes at USNO Flagstaff. Residuals are from a direct star-to-star comparison between May 11 and May 12, 2008.
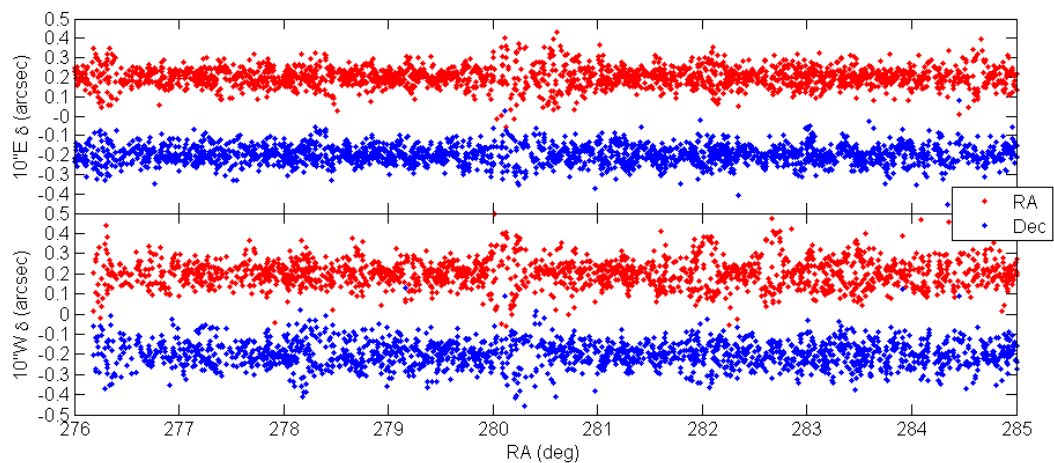
Figure 66. Spline-subtracted residual noise for the two UNMCO telescopes. Residuals are from direct stellar comparisons between the nights of May 13 and May 18, 2009.

143
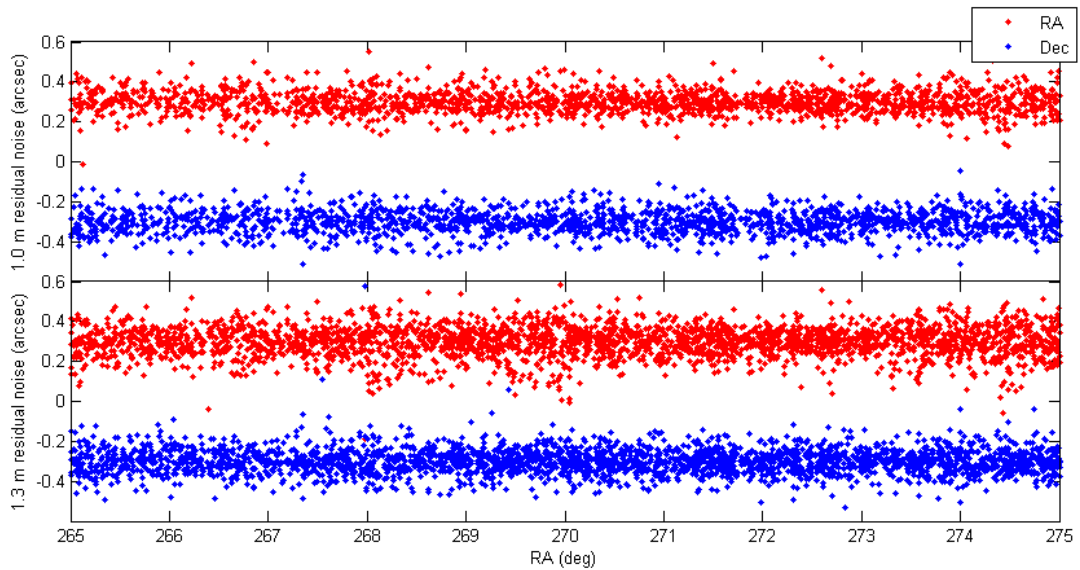
Figure 67. Spline-subtracted residual noise for both the 1.0 meter and 1.3 meter telescopes at USNO Flagstaff. Residuals are from direct star-to-star comparison between May 11, 2008 data and CMC14 positions. Note difference in y-scales between this and the previous figure.

After extensive analysis of image conditions ranging from external temperature to stellar FWHM, the source of the elevated noise was attributed to variations in the telescope plate scale with time and with y-position. Where plate scales are the same between matched observations, the noise levels in the residuals are small; however, mismatches in plate scales will result in increased noise levels in the residuals. The row blow-ups can be attributed to plate scale differences that vary both with time and y-position on the plate (*e.g.* at a given time a fraction of the frame will be skewed into a trapezoid). Stars crossing the field at a particular time will vary in x-position with their y-position on the chip Column blow-ups result from vertical scale changes that cause the range of declinations viewed by the focal plane to differ between nights and telescopes. Because we are operating in drift-scan mode, if a scale mismatch continues uniformly for more than a frame-length, we would expect the errors in row position to average out, but

144

column errors to remain (as appears in Figure 68, where a constant plate scale difference causes a continuous spread in Dec prior to correction). This sort of noise resulting from fluctuations in plate scale would be expected to also appear in comparisons to external standard catalog positions, such as CMC14, however, the effect should be considerably damped as only the image positions will experience plate scale changes.

Correcting the blow-ups is a straightforward matter of fitting the plate scale of one night to that of the other on frame-length or (preferably) shorter timescales. For the direct night to night comparison used with the June 2007 data, a linear fit to the column positions of a set of stars as observed one night to the same set the other night on a half-frame basis is used to correct the column blow-ups. Row blow-ups are corrected with a fit of the row positions of one night to the corresponding row and column positions of the next night with row scale set to one. It is important to note here that changing plate scale in the row direction almost certainly affects the images and would manifest as offsets in row position as a function of time (for scale differences varying on sub-integration timescales), not unlike anomalous refraction. There is, however, no possibility that anomalous refraction is an artifact of fluctuating plate scales considering that AR equally affects both row and column positions and a plate scale mechanism for systematic column displacements does not exist. A subset of the June 2007 residuals corrected for plate scale changes as described is shown in Figure 68. The comparison with the uncorrected residuals not only provides a marked demonstration of the efficacy of the solution, but indicates that contrary to being rarer than row blow-ups, the seeming lack of

column blow-ups is due to the column residuals being in a nearly constant state of blow-up and not undergoing the isolated expansions of the row residuals.
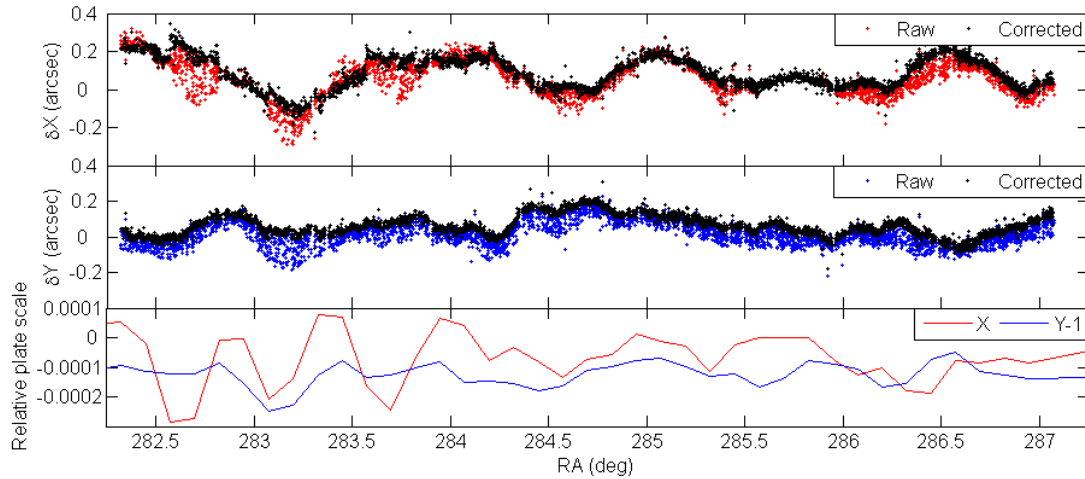
Figure 68. X and Y residuals resulting from the direct night to night comparison of 11 June 2007 to 13 June 2007 before and after correcting for the differences in plate scale shown in the third panel.

In light of eliminating or understanding all potential sources of error and causes of AR or AR-like effects, an important test was verifying that the catalogs to which we compared our data were not themselves contributing to the motions of the residuals. A comparison of residuals referenced to UCAC2 and those referenced to CMC14 shows a very high degree of correlation, with any differences being far outweighed by anomalous refraction in the images (Figure 69). This comparison is an excellent demonstration that fluctuations we are seeing in the residuals are not due to the catalog to which we compare the positions and shows that we can interchangeably use either the UCAC2 or CMC14 catalogs without changing the results.

146
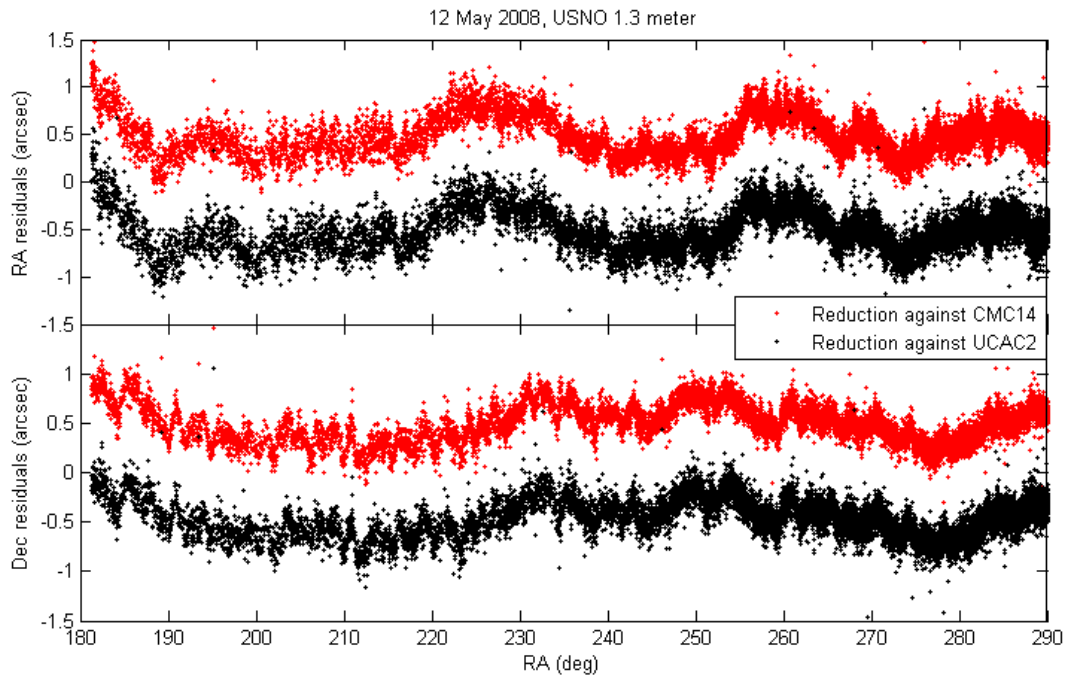
Figure 69. Comparison of residuals referenced to UCAC2 and to CMC14 for one night of 1.3-m observations.

147

4.2 Atmospheric observations

Throughout the observing program collocated surface atmospheric observations were made simultaneously with the astrometric observations in order to allow better understanding of both the weather conditions associated with the occurrence of AR and to relate AR to the pressure disturbances caused by its hypothesized source, AGWs. The goals of the atmospheric observations were first to observe AGWs while making astrometric observations to see whether AGWs are present when AR is seen in the data. Should that be the case, we then sought to examine the characteristic relationship between the two. The main atmospheric instrument was an array of microbarograph pressure sensors (see Section 4.2.1) designed to detect boundary-layer AGWs. We also employed Davis Instruments Vantage Pro 2 weather stations (see Section 4.2.2) during all observations to measure wind speeds and directions, as well as other surface atmosphere characteristics such as relative humidity and temperature. With the weather station we desired to determine primarily if AR had any dependence on wind speed and direction, possibly stemming from wave generation by airflow over prominent topographic features. ALE, the Astronomical Lidar for Extinction (see Section 4.2.3) was operated concurrently with several of the UNMCO observations to look for any indication of wave-like dynamics in the aerosol structure of the boundary layer.

4.2.1 Microbarographs

Microbarographs are passive, remote surface-based atmospheric sensors that have been employed in boundary layer AGW studies for over 50 years (*e.g.* Gossard & Munk 1954). A microbarograph is a highly sensitive pressure sensor, capable of detecting differential surface pressures on the order 1-100 microbars (μb). While microbarographs

can come in a variety of forms, the most common design employs a piezoelectric diaphragm which is mounted between a reference pressure chamber and the atmosphere. The reference chamber is set to ambient atmospheric pressure at the start of observations and small fluctuations in pressure from this initial value cause the diaphragm to distort. When bent or compressed, piezoelectric materials will produce a voltage proportional to the degree of strain (and likewise will bend when a voltage is applied to them), providing a means to record very subtle atmospheric pressure changes (Nappo 2002). The reference volume of the microbarograph is configured with a slow leak or reset to atmospheric pressure at standard intervals to prevent overpressure. We employ Omega Engineering PX278 low pressure transducers for this central element in the array. These differential pressure transducers (DPTs) are sensitive to three selectable differential pressure ranges full scales of ± 0.31125 mb, ± 0.6225 mb and ± 1.245 mb with a measurement accuracy of 0.1% of the full scale or 6 µb in the smallest range setting.

When an atmospheric gravity wave in the stable boundary layer passes over a point on the surface, it causes a small change in the surface pressure at that point because of the change in height of the air column. Quantitatively, the relationship between the wave displacement, ζ, at an altitude, z, and the pressure perturbation at the ground is,

$$P_0 = \zeta \frac{\rho_s \left(\omega/k \right)^2}{z} \left(\frac{n_1 z}{\sin(n_1 z)} \right), \tag{35}$$

where $n_1 = k|(N/\omega)^2 - 1|^{1/2}$, N is the Brunt-Väisälä frequency in the surface layer, and ρ_s is the atmospheric density at the ground (Gossard & Sweezy 1974). The phase speed (ω/k) is measured with respect to the mean background wind and is related to the

perturbation pressure and perturbation wind velocity, u, during positive pressure perturbation through the impedance relationship (Gossard & Munk 1954),

$$\frac{\omega}{k} = \frac{1}{\rho_s}\frac{P_0}{u}.$$ (36)

The pressure differences recorded by a microbarograph are the sum of all pressure effects occurring in the column of air directly above that particular sensor within the boundary layer.

A single point source microbarograph will see a superposition of boundary-layer AGWs in the column of air directly above the sensor. From this it is possible to measure relative wave amplitude and frequency with respect to the surface. The addition of wind speed and direction measurements at the surface allows calculation of the wave phase velocity using Equation (36). Creating an array of multiple microbarographs dramatically increases the utility of the instrument. With an array of three or more microbarographs, it becomes possible to not only detect the presence of AGWs through their effects on surface pressure, but also to determine wavelength, phase speed, direction of travel, period and amplitude.

The MicroBarograph for Anomalous Refraction (μBAR) being used for this project is a variation on the standard microbarograph design. Rather than measure the atmospheric pressure at a point relative to a reference volume, we directly measure the pressure difference between two points a set distance apart. The primary information we require from μBAR to relate AGWs to anomalous refraction is the pressure gradient across the field of view of the telescope. An AGW passing over the telescope causes the

atmospheric strata at wave altitude to tilt, hypothetically resulting in the anomalous refraction seen by the telescope on the ground. This tilt is translated into an atmospheric pressure gradient at the surface that should be measured by the differential microbarograph. Removing the reference volumes and measuring the differential pressure between two points gives us this desired pressure gradient directly and removes any concern of maintaining uniform reference pressures, or the need to periodically reset the pressure of the reference volume to prevent overpressurization. The direct differential method also has the potential for observing standing waves (which would register as a constant pressure difference between the two points), something that cannot be measured by traditional (point source) microbarographs because of the regular resetting of the reference volume to prevent overpressurization.

The primary downfalls of the differential microbarograph array are its inability to measure the absolute pressure amplitude of a wave (*i.e.* the pressure perturbation relative to atmospheric pressure) and that its pressure range must extend to significantly lower pressures than those observable by a point source microbarograph. (While the absolute surface pressure perturbation of an AGW may be on the order of 100 microbars, the pressure difference between two points on the wave separated by ten meters is proportional to the amplitude divided by the wavelength and may only be 1 – 10 microbars or less.) In order to obtain the broadest possible spectrum of information with the microbarograph array we employ a combination of several differential microbarographs in tandem with one traditional point source microbarograph.

151

Extensive tests were performed on the microbarograph array prior to deployment in the field to devise the most robust instrumentation and to fully characterize the instrumental response under various operational conditions and atmospheric inputs.

Initial testing centered on the DPTs without any net pressure inputs. To measure the electrical noise characteristics of the DPTs, we operated the instruments in a controlled thermal environment with and without electrical shielding. The unshielded DPTs were found to experience significant radio frequency interference at 60 Hz, 120 Hz and multiple harmonics thereof. Placing the sensors in a shielded environment (both a solid metal enclosure and a wire mesh enclosure were tested) eliminated nearly all of the observed noise and resulted in a white noise spectrum from the unpressured sensors (Figure 70).
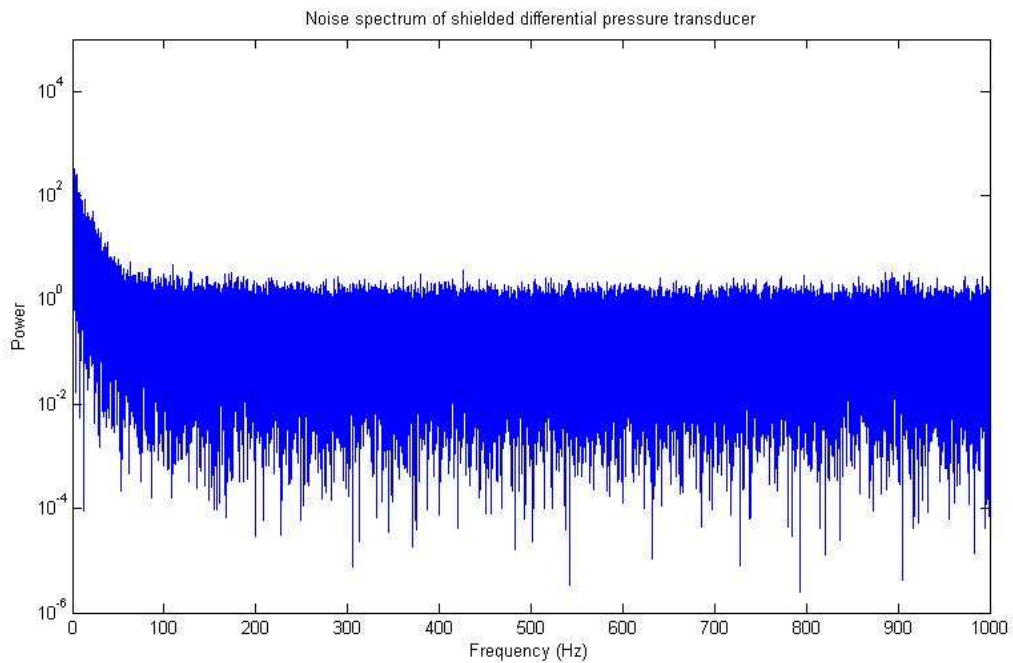
Figure 70. Spectrum of shielded differential pressure transducer with no net pressure input.

As an additional means of reducing noise at very high frequencies we added an RC-circuit low-pass electrical filter (Figure 71). The RC-circuit is composed of a capacitor wired in parallel and a resistor in series with the signal path from the pressure sensor to the data acquisition device. Because the impedance of a capacitor decreases with increasing frequency, high frequency signals will be mostly shorted out by the capacitor with any remaining signal attenuated by the resistor. The cut-off frequency of an RC circuit with resistance, R, and capacitance, C, is, $f_c = 1/2\pi RC$ and is defined as the frequency at which the capacitive reactance in ohms equals the resistance and above which the filtered signal falls below 70.7% of the input signal. For the electric filters used in μBAR, we found a resistance of 30 kilo-ohms and a capacitance of 20 micro-Farads provided the best damping of high frequency noise. The resulting filtered signals (and comparison to the unfiltered signal) are shown in Figure 72.
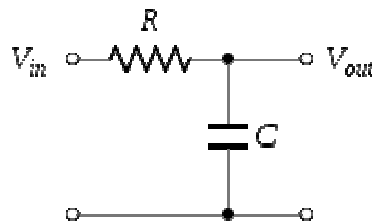
Figure 71. RC low-pass electrical filter circuit diagram.

A second test was devised to determine the effect of the thermal environment on the DPT pressure reading. This was measured by operating the sensors with the opposite differential ports connected such that there is zero differential pressure input under a wide range of external temperatures. The published temperature error (or change in zero point pressure with temperature) of the particular DPTs used is ±0.02% of full scale per degree Celsius, which for these tests (full scale of 1.245 mb) amounted to 0.25 μb/°C. Prior to

the second phase of observations in 2008 several additional DPTs were purchased from Omega to allow the expansion of the μBAR to three differential microbarographs and one point source microbarograph. Two of these sensors have a new circuit board design which is considerably different from the original DPTs. This change in circuit boards was not accompanied by a change in published instrument specifications. The actual observed temperature dependences for three DPTs are plotted in Figure 73 and averaged about 0.24 μb/°C for the two original design sensors and about 4.6 μb/°C for the new sensor. The new-type sensors also exhibit significant hysteresis between the temperature dependence for increasing temperature and that for decreasing temperature (which causes the large spread in the error bars of the new style temperature trend). It is not clear why the newer sensors exhibit stronger temperature dependences than the old sensors; however, this effect is corrected during data processing.
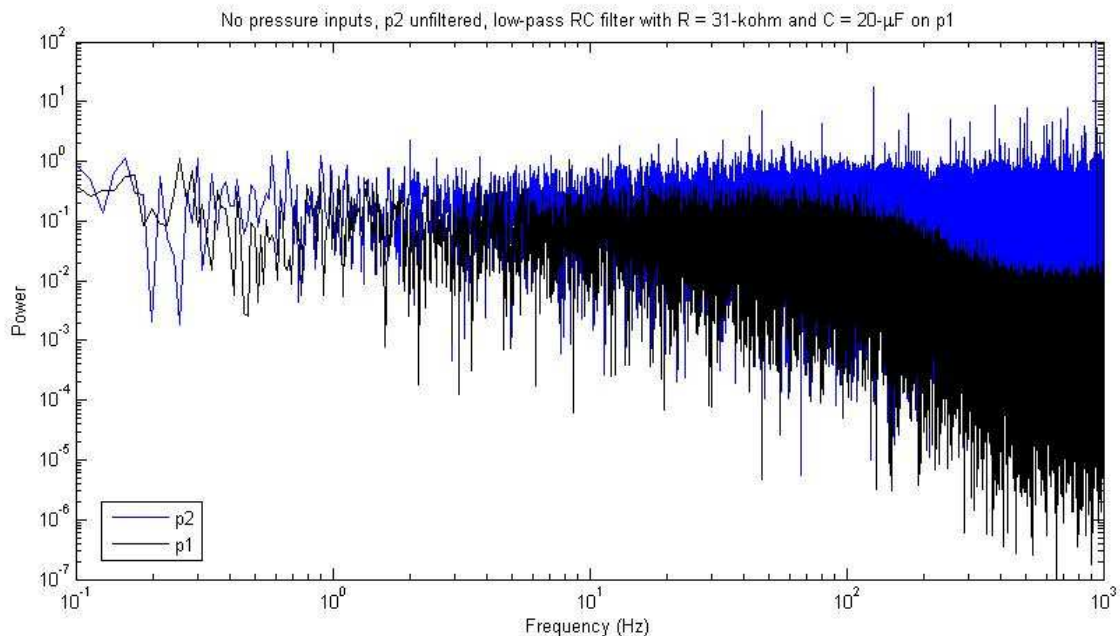
Figure 72. Frequency output of two pressure sensors with zero net pressure input. P1 has been filtered using an RC low-pass circuit while p2 has been left unfiltered as a reference.

During actual field operation of the sensors, temperature readings taken by thermistors in the same enclosure as the DPTs are included in the data and the above relationships are applied to correct for temperature induced errors in the pressure readings (DC offsets apparent in Figure 73 are also accounted for).
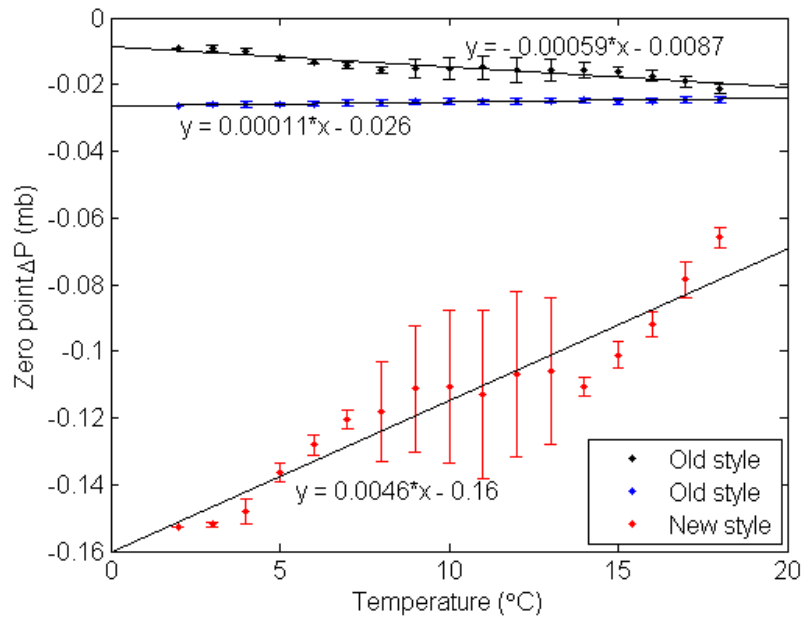
Figure 73. Temperature error of DPTs operated with zero differential pressure input. Pressures are recorded at 100 Hz and averaged into one degree Kelvin temperature bins with standard deviations indicated by the error bars.

One of the greatest challenges inherent in the design of a microbarograph intended for any purpose is the filtering of atmospheric noise to isolate the signals of interest. The primary component of this noise for nearly all microbarographs regardless of intended use is wind induced pressure fluctuations. These fluctuations are highly random, occurring on timescales of seconds and having very large amplitudes, of order hundreds of microbars to millibars – often much larger than the signals one is trying to observe. The requirement of such a filter is that it significantly damps (by two or three orders of

155

magnitude) the noise fluctuations while allowing the signals of interest to reach the sensor unchanged.

Atmospheric science literature is replete with articles (*e.g.* Hedlin & Alcoverro 2005, Hedlin *et al.* 2000, Daniels 1958) extolling the merits of various filter designs ranging from the very simple to the incredibly elaborate. By far the most popular design in current microbarograph construction is the rosette (Alcoverro & Le Pichon 2005), an extensive array of pipe inlets radiating out from the pressure sensor. Small diameter pipes are arrayed along six to eight radials extending several tens of meters out from the pressure sensor inlet. At the end of each radial pipe, another radial array of pipes extends another few tens of meters around the inlet. The full filter is anywhere from fifty to two hundred meters in diameter and works by mechanically summing all pressure fluctuations occurring over the area of the filter such that only isolated disturbances and large scale systematic phenomena are detected.

This filter design is primarily used by nuclear test ban treaty (NTBT) monitoring microbarographs (Alcoverro & Le Pichon 2005). The signals of interest for these sensors are short timescale, very isolated infrasonic pressure fluctuations resulting from tests of nuclear weapons. Because these signals travel at the speed of sound and radiate in all directions from the source (as opposed to AGW signals which propagate primarily vertically from the disturbance), they effectively encounter all input ports of the rosette filter simultaneously and thus are fully transmitted to the pressure sensor.

156

The requirements for a filter for μBAR are significantly different from those for a NTBT microbarograph. The pressure disturbances we wish to measure are coherent over spatial scales of tens of meters to kilometers and are very slowly varying, not isolated and short timescale like the pressure wave from a nuclear blast. Recall also that the μBAR is a differential microbarograph as opposed to a point source microbarograph and we are measuring the difference in atmospheric pressure between two points separated by ten meters. This second characteristic in itself rules out the use of many standard microbarograph filter designs, including the rosette. We require that the pressure observed by each input port of the differential sensor be summed over an area of radius much smaller than the separation of the two ports, otherwise the observed pressure difference is significantly diminished.

An alternative to the rosette filter is the wind fence (Hedlin & Raspet 2003). The wind fence is built several feet high surrounding the microbarograph inputs. A number of designs exist, however the general principle of all wind fences is that turbulent eddies passing through the openings in the fence (a typical design has 50% porosity, either by a mesh screen, wood slats, etc.) are broken up resulting in laminar air flow some distance behind the fence. Based on studies of the airflow behind a wind fence (Hedlin & Raspet 2003), in order for the airflow passing over the inlet ports of the microbarograph to be laminar, the fence would have to surround the ports at a radius of tens of meters. The scale of these installations is not practical for this stage of the research, particularly given that the microbarographs will be short term temporary installations at the observing sites. We would require an extensive wind fence network as the μBAR array itself is ten meters

in diameter. In most observatory situations, a wind fence enclosing an area tens of meters in diameter is not a practical installation.

Extensive research and testing went into the design of the filters used on the μBAR. Early designs ranged from baffles based on mufflers or silencers to porous inlet enclosures (such as foam spheres) designed to average pressure variations over a surface. None of these designs were effective at reducing high frequency noise by much more than a factor of 10. Two of the more notable designs tested were the sand filter and the foam block filter

The sand filter involved burying each inlet port under several inches of sand. This was based on the microbarographs used in an Antarctic AGW research project (Anderson *et al.* 1992). Anderson *et al.* (1992) found that burying the microbarograph inlets under a meter of snow provided very effective high frequency filtering. Burying the inlets under a few inches of sand proved less effective than expected (Figure 74a). The foam block filter was based on the porous foam spheres commonly used to filter out wind noise for microphones in acoustic sound systems, and has been tested for use in microbarographs (Hedlin & Raspet 2003). The inlet ports were inserted into the center of the foam block, which then acted to average all pressure fluctuations over the surface of the block. The filtering efficiency of this setup is shown in Figure 74b. The foam block filter construction was used on the μBAR during the first observing run at the USNO Flagstaff in June of 2007.
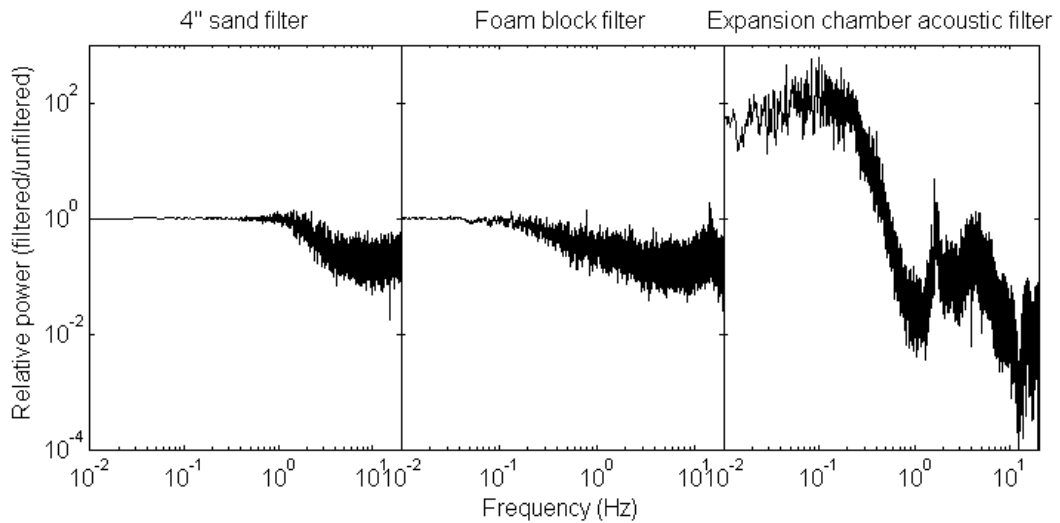
Figure 74. Comparison of acoustic filter designs tested for use with µBAR. a) The left plot shows the relative filtering efficiency of burying the pressure inlet under four inches of sand. b) The center plot shows the relative filtering efficiency of embedding the pressure inlet in an 8" foam block. c) The right plot shows the filtering efficiency of the three part expansion chamber acoustic low-pass filter used in the final microbarograph design.

Taking a different approach to designing the filters, we were finally able to land on a highly efficient filter that meets the requirements of µBAR. The filter design stems from basic acoustic theory, particularly the design of an acoustic low-pass filter employing an expansion chamber (Figure 75). The expansion chamber operates by using the change of impedance where the inlet pipe changes in diameter to selectively reflect waves of high frequencies.
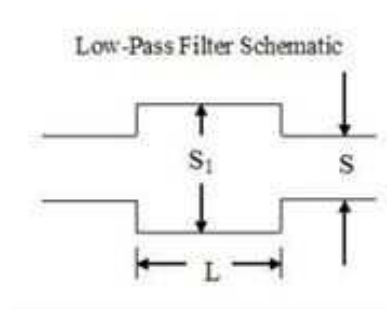
Figure 75. Basic low-pass acoustic filter.

Quantitatively this expansion chamber filter is acoustically equivalent to an infinite pipe of impedance $Z_0 = \rho_0 c/S$, where ρ_0 is the density of air in the pipe, c is the speed of sound and S is the cross-sectional area of the pipe, with a side branch located at $x = 0$ of impedance $Z_b = R_b + iX_b$. R_b and X_b are the resistance and reactance respectively. The incident and reflected pressures at the branch are $p_i = Ae^{i\omega t}$ and $p_r = Be^{i\omega t}$ respectively. Continuity of pressure tells us that at the boundary where the side branch enters the infinite pipe, $p_i + p_r = p_1 = p_2$, where p_1 is the pressure in the side branch and p_2 is the pressure in the infinite pipe beyond the branch. Additionally, continuity of volume velocity, U, requires that the total volume velocity remain constant before and after the junction (Bernoulli's principle): $U_i + U_r = U_1 + U_2$. The volume velocity is defined as $U = p/Z$. Dividing the two continuity equalities gives $(U_i + U_r)/(p_i + p_r) = U_1/p_1 + U_2/p_2$. Substituting the values of U and p into this equation and solving for A and B gives us $B/A = (-\rho_0 c/2S)/(Z_b + \rho_0 c/2S)$. The reflection coefficient at the junction is defined as the squared modulus of the reflected amplitude divided by the incident amplitude,

$$\mathcal{R} \equiv |B/A|^2 = \left(\rho_0 c/2S\right)^2 / \left(\left(R_b + \rho_0 c/2S\right)^2 + X_b^2\right). \tag{37}$$

The impedance of the side branch is a pure reactance, *i.e.* $R_b = 0$. Reactance is qualitatively the restoring force in a periodic motion.

Using an AGW as a mechanical analogy, the resistance is obviously the air resistance encountered by a mass moving at a nonzero airspeed and depends only on velocity, while the reactance is the downward force of gravity at the crest of the wave or the upward force of buoyancy at the trough of the wave and depends on the wave phase. Reactance

is quantitatively defined as $X_b = \omega m - 1/C\omega$, where m is the intertance (acoustic equivalent of electrical resistance, which in this case is zero) and $C = S_1 L / \rho_0 c^2$ is the compliance (equivalent to capacitance). Because the transmitted power is equal to the total incident power minus the reflected power, the fraction of power transmitted is, $\mathcal{T} = 1 - \mathcal{R}$. From the above equations we can show that,

$$\mathcal{T} = \frac{1}{1 + \left(\frac{S_1}{2S} kL\right)^2},$$ (38)

where $k \equiv \omega/c$ is the wavenumber of the acoustic wave. This value of the transmission coefficient is valid for values of $kL < 1$.

For values of $kL > 1$ we need to consider the effects of the impedance change at each boundary of the expansion chamber-infinite pipe system on the incoming, reflected and transmitted waves. We consider a system of three consecutive regions (Figure 76) with impedances Z_1, Z_2 and Z_3 respectively with the boundary between the first two at $x = 0$ and the boundary between the second two at $x = L$, where L is the length of the expansion chamber (see Figure 75). In the first region the incoming and reflected waves are described by $P_i = p_i e^{i(\omega t - k_1 x)}$ and $P_r = p_r e^{i(\omega t + k_1 x)}$ respectively. Likewise in region 2, $P_a = A e^{i(\omega t - k_2 x)}$ and $P_b = B e^{i(\omega t + k_2 x)}$, and in region 3, $P_t = p_t e^{i(\omega t - k_3 x)}$. We can apply the continuity of pressure and volume velocity at the boundaries of $x = 0$ and $x = L$ in a similar fashion as before to get $Z_1(p_i + p_r)/Z_2(p_i - p_r) = (A + B)/(A - B)$ at $x = 0$ and $\left(A e^{-ik_2 L} + B e^{ik_2 L}\right)/\left(A e^{-ik_2 L} - B e^{ik_2 L}\right) = Z_3/Z_2$ at $x = L$. After some extensive

algebra we find, $\frac{p_r}{p_i} = \frac{\left(\left(1 - \frac{Z_1}{Z_2}\right)\cos k_2 L + i\left(\frac{Z_2}{Z_3} - \frac{Z_1}{Z_2}\right)\sin k_2 L\right)}{\left(\left(1 + \frac{Z_1}{Z_3}\right)\cos k_2 L + i\left(\frac{Z_2}{Z_3} + \frac{Z_1}{Z_2}\right)\sin k_2 L\right)}$. Recognizing that for the expansion

chamber filter, the pipe on either side of the chamber (*i.e.* regions 1 and 3) has the same diameter and therefore the same impedance and noting that $Z_3/Z_1 = S_1/S$, we find for the transmission coefficient,

$$T = 1 - \left|\frac{p_r}{p_i}\right|^2 = \frac{4}{4\cos^2 k_2 L + \left(\frac{S_1}{S} - \frac{S}{S_1}\right)^2 \sin^2 k_2 L}.$$ (39)
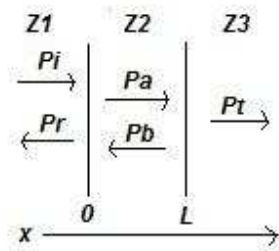
Figure 76. Schematic of the boundary value problem consisting of a wave passing through three consecutive regions with different impedances.

From Equations (38) and (39) it is clear that the transmitted acoustic power of an expansion chamber depends primarily on the relative cross-sections of the inlet/outlet pipe and expansion chamber as well as the length of the chamber itself. At low frequencies, changing the cross-section or length of the expansion chamber by an equal amount will have the same effect on the transmission. However, at high frequencies (*kL*>1), changing the length of the chamber leaves the maximum filtering efficiency unchanged, but changes the frequency characteristics of the filter, while changing the ratio of cross-sections increases the efficiency at all frequencies (Figure 77). The periodic nature of the transmission coefficient at high frequencies means that a single transmission chamber tuned to attenuate moderate frequency noise will still allow a comb of higher frequencies to pass unfiltered. Employing a series of transmission chambers of carefully chosen varying lengths can obviate this effect.
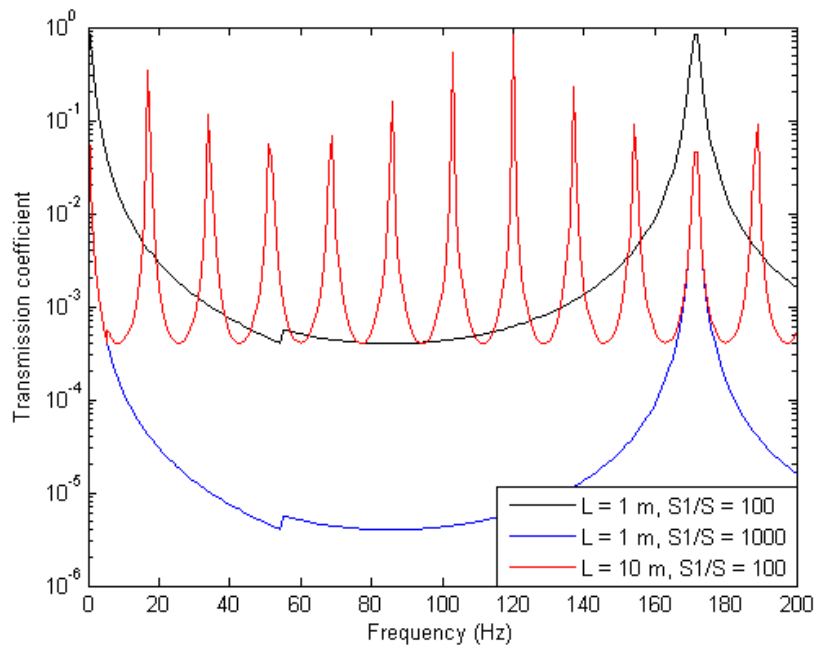
Figure 77. Transmission coefficient of an expansion chamber low-pass acoustic filter and the effect of changing either the relative cross-sections of the chamber and pipes or the length of the chamber.

The expansion chamber filter for µBAR was engineered to attenuate frequencies above 0.05 Hz while allowing frequencies of 0.02 Hz and lower, typical values for AGWs, to pass unaffected. To achieve these ends the acoustic filter employs a series of three expansion chambers installed in each leg of the microbarograph. Standard four inch and one inch PVC pipe was used for the chambers (S1) while one inch disks with 0.062 inch holes served as the connecting pipes (S). The first chamber was a five foot section of four inch diameter PVC with one end closed by a roughly half inch thick cap with a 0.062 inch hole opening to the atmosphere. The opposite end is connected to a two foot section of four inch diameter PVC with a disk as described above separating the two chambers. The third chamber is a ten foot long section of one inch PVC connected to the two foot chamber with another 0.062 inch hole. The theoretical transmission of each element and

163

the combined transmission of the complete filter are shown in Figure 78. The high frequency filtering characteristics are clearly sufficient for our purposes, where a reduction of three orders of magnitude would be acceptable. For the low frequencies, the transmission is reduced by more than an order of magnitude by 0.05 Hz but is only minimally attenuated below 0.02 Hz as we require.

The actual filtering efficiency of the expansion chamber system was tested by running two differential microbarographs side-by-side for a full night with one filtered and one unfiltered. The result of this test is shown in the third panel of Figure 74. The overall efficiency of the filter is clearly better than any of the previous designs, but does not share much in common with the theoretical efficiency in Figure 78. The discrepancy can be traced to several elements of the design. The first consideration is that the theoretical transmission coefficient assumes that each expansion chamber is inserted into an infinite pipe, which is clearly not the case in this instrument, where the "pipes" connecting the chambers are not more than an inch long. Additionally, the last expansion chamber in this setup is connected to 3/16 inch tubing rather than the 0.062 inch "pipe" used between the chambers. The enhanced sensitivity at very low frequencies likely results from impedance differences between the open-ended one-inch pipe (used for the non-filtered case) and the expansion chamber system. The final thing to consider is that although the filter was built as carefully and precisely as possible, it cannot be guaranteed that there are no leaks in the system, the presence of which would reduce the filter efficiency. During the operation of the microbarograph extensive efforts were made to reduce this

last variable by sealing all threaded connections with Teflon tape designed for gas pipe connections or Teflon pipe joint compound.
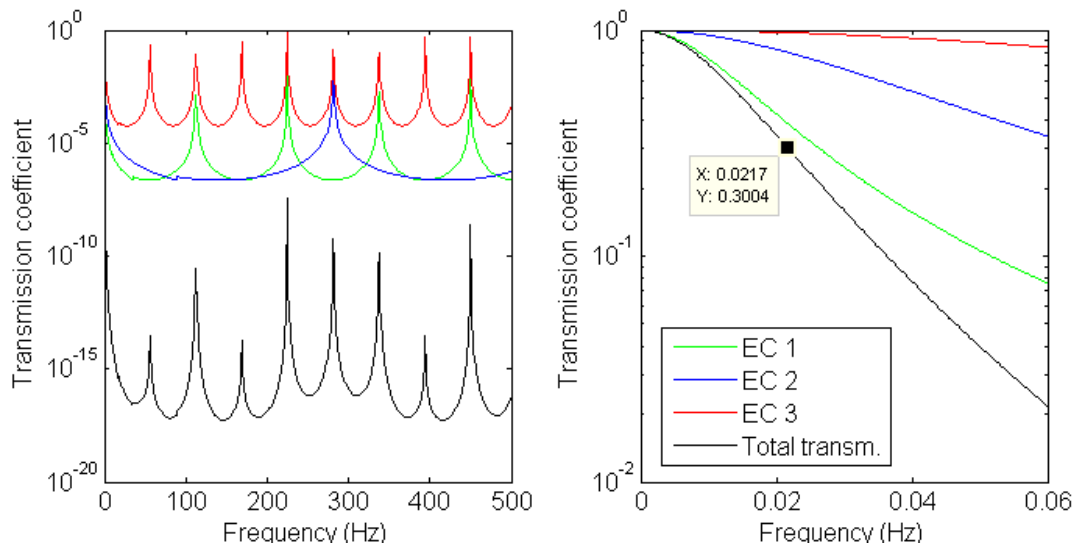
Figure 78. Transmission coefficients of each expansion chamber (EC) in the μBAR acoustic filter when operated in isolation and total transmission of the combined filter system at high frequencies (left) and low frequencies (right). Total transmission is reduced by 3 dB at just over 0.02 Hz.

A final part of the microbarograph design was a pressure calibration chamber initially intended to absolutely calibrate the DPTs with a known input pressure. By measuring the individual response of the DPTs when a known identical pressure difference was applied to each sensor, any discrepancies in the response as a function of pressure input could be accounted for.

The calibration device was designed based on the ideal gas law – at constant temperature and number of moles of gas, the change in pressure of a system is directly proportional to the change in volume of the system. The design consists of two identical cylindrical chambers of precisely known volume in direct thermal contact (the cylindrical volumes are stacked end on end, separated by one inch of aluminum), with the volume of one

chamber variable by means of a piston. The entire system was constructed out of aluminum with the intention of maximizing heat transfer to and from the system such that near-isothermal conditions could be maintained. All openings from the chambers to the atmosphere are sealed with o-rings (*e.g.* the lid of the top cylinder, the piston, *etc.*). The dimensions of the system are as listed in Table 1 and the system itself is shown in Figure 79.

Chamber radius	0.05715 m
Chamber length	0.08255 m
Chamber volume	8.47×10^{-4} m^3
Piston radius	0.00238 m
Piston travel	± 0.00005 m - ± 0.005 m
Connecting tubing volume	2.766×10^{-5} m^3

Table 1. Dimensions of the microbarograph pressure calibration chamber.

Figure 79. Pressure calibration chamber for μBAR. On the right are the stacked chambers, labeled "H" and "L", with the top chamber opened to show the interior volume. The lid of the top chamber is on the left with the piston hole in the center and the piston sitting on top. The vertical posts on the lid are for mounting the stepper motor and the "wings" on the piston are guides to keep the piston from rotating.

The piston is controlled using an Anaheim Automation stepper motor with a step resolution of 1/200th of a revolution and accuracy of two steps, and a lead screw with 45 threads per inch. The resulting linear piston motion resolution is 2.8 microns with an accuracy of five microns. The motor is controlled through serial port communication with a motor controller interfaced through a LABVIEW program which also queries the controller encoders at roughly one Hertz to log the piston positions.

The dimensions of the calibration system allow isothermal pressure differences between the two chambers from approximately 10-μb to 160-μb. To calculate the actual isothermal pressure difference between the two chambers following a volume change in one chamber, we need to know the number of moles of gas in one chamber, the temperature of the whole system, and the volume change. The number of moles of gas can be determined from $N = \rho V / M$, where the density of air, $\rho = P_d M_d / RT + P_v M_v / RT$, is a function of the pressure and molar mass of dry air (subscript d) and water vapor (subscript v), and V and M are the volume of the chamber and mean molar mass of air respectively. The pressure difference is then,

$$\Delta P = \frac{NkT}{(1 + CTE(Al)(T - 293))^3} \left[\frac{1}{V} - \frac{1}{V - \Delta V} \right], \tag{40}$$

where *CTE(Al)* is the coefficient of thermal expansion of aluminum (with a published value of 23.1×10^{-6} m/m/K). This takes into account the temperature of the system (assuming equilibrium with the atmosphere) and the thermal expansion/contraction of the aluminum.

Prior to operating the calibration system, both chambers are briefly opened to the atmosphere to remove any residual pressure differences and N is calculated based on atmospheric conditions (pressure, temperature and relative humidity) read by a weather station. The calibration sequence involves first moving the piston in several steps to a maximum extension of 0.005 meters, returning to the zero point, and then withdrawing the piston in several steps to a minimum position of -0.005 meters, resulting in maximum pressure differentials of a little more than ±0.1 mb. An example of the calculated input pressures and pressure differences measured by the pressure sensors during a calibration sequence is shown in Figure 80.
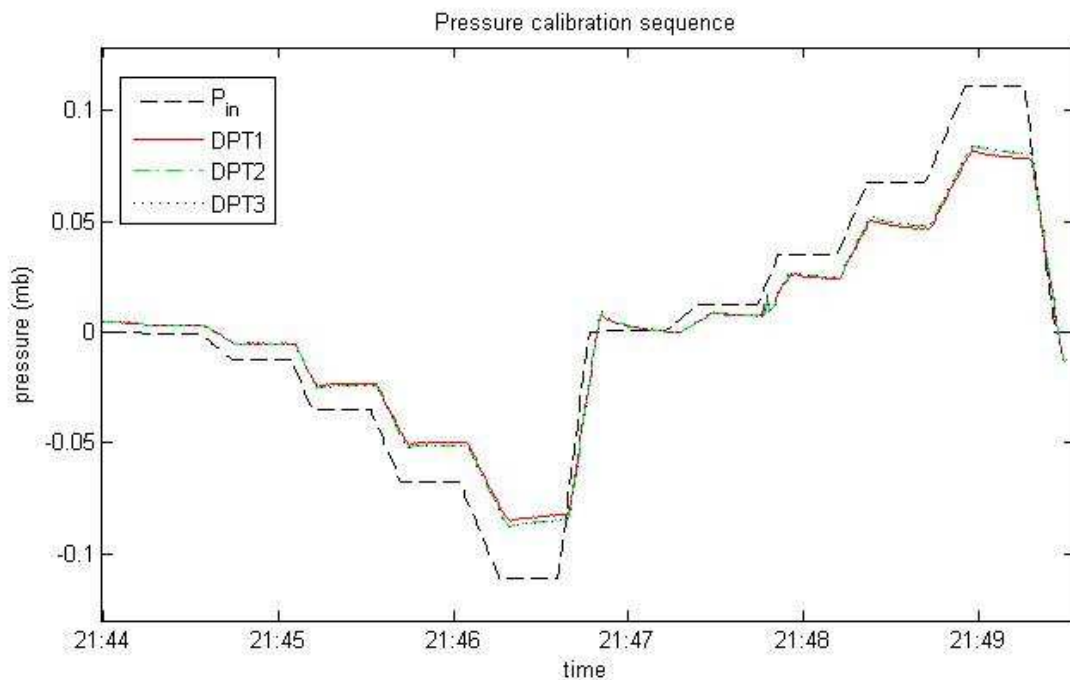
Figure 80. Calculated (P_{in}) and measured (DPT#) pressure differentials during a pressure calibration sequence.

It is clear from Figure 80 that there are systematic differences between the calculated input pressure and that recorded by the three pressure sensors. These offsets are nearly

identical in all three sensors, and appear to be roughly proportional in scale to the calculated input pressure. The pressure measured by DPT1 as a function of the calculated input pressure is plotted in Figure 81. The pressure sensors show an approximately linear response to the input pressure with a constant of proportionality of approximately 0.73.
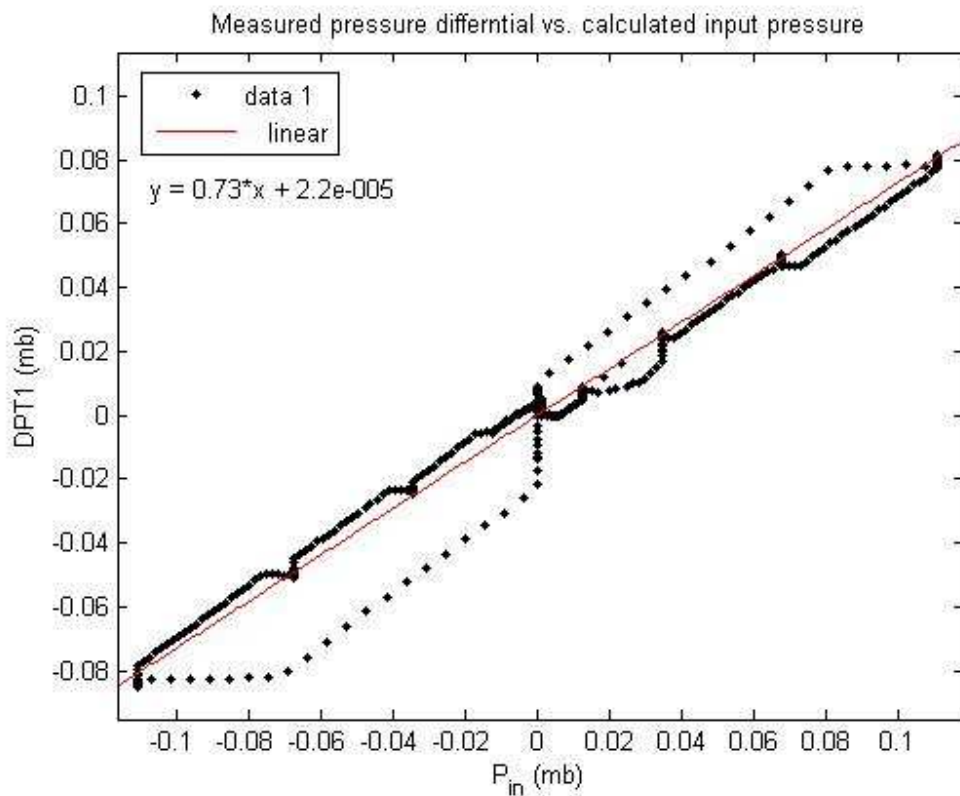
Figure 81. Differential pressure measured by DPT1 as a function of the calculated input pressure during a pressure calibration sequence.

Despite attempts to reduce system variables, including leak mitigation and temperature regulation, there remain a number of variables which may be contributing to the disparity between the calculated and observed pressures. The issue of leaks in the system was encountered early on in the design and construction process. While a large number of

measures, including compression fittings and o-ring seals were put in place to prevent leakage, due to the miniscule pressure differences involved (pressures in the system can be considered equal to the outside atmosphere), actually detecting or ruling out leaks to the microbar level is not practicable.

In the process of attempting to limit leaks in the system, we discovered that what was originally considered to be large scale leakage was actually uneven heating of the tubing connecting the pressure calibration chamber to the DPTs. The primary source of the heating was determined to be the solenoid valves used to switch between the calibration tubing configuration and the atmospheric measurement configuration. Once the solenoids were removed, pressure variations in the closed system dropped by an order of magnitude. The remaining variations were on the order of microbars per minute and could not be reduced further. These are likely due to a combination of slow leaks and small scale temperature differences. While these small variations are evident in the calibration data, they cannot be the primary cause of the observed disparity during the calibration sequence (due to the differences in the timescales involved).

One source of the temperature differences in the system which we were able to control was non-uniformity in the volume of the tubing connecting the pressure sensors to the calibration system and atmosphere. For tubing with an inner diameter of 3/32 inch, if the two ports of a DPT were attached to separate closed sections of tubing with lengths of 19 inches and 20 inches initially having the same pressure, the pressure difference between the ports due to a temperature change of one Kelvin is calculated to be approximately two

170

microbars (assuming negligible tube expansion). The lengths of tubing used to connect the sensors was generally longer than 20 inches and for the science observations the tubes are not closed, so the effect of non-uniform tube lengths would be minimal. However, during calibration, the system is airtight so non-uniform pressure changes due to varying tube volumes were a far more significant concern and extensive pains were taken prior to the 2008 NOFS observing runs to connect all sensors with identical lengths of tubing

The most likely source of the proportional offsets is the fact that the input pressures are calculated assuming an isothermal system, when in reality the system lies somewhere between isothermal and adiabatic. If we examine an example of the pressure calibration system operated at sea level with a 0.005 m piston displacement, we can determine the corresponding pressure change for an isothermal vs. adiabatic system. If the system is isothermal, the pressure is calculated from Equation (40). Assuming for this example that the system is at room temperature, we find a pressure difference of 155 microbars. For an adiabatic system, $PV^\gamma = const$, where $\gamma = C_p/C_v = 1.4$ for air (C_p and C_v are the heat capacities of air at constant pressure and volume respectively). Using standard sea level pressure and the initial volume (piston at zero) of the system, we calculate the constant to be 5.067. We then calculate the adiabatic pressure change corresponding to a 0.005 m piston displacement to be 68.3 microbars. A similar calculation performed for a piston displacement of 50 microns shows isothermal and adiabatic pressure changes of 7.3 microbars and 1.1 microbars respectively. Thus for a given volume change, the range in which the resulting pressure change may lie depending on the thermodynamic nature of the system is large compared to the actual pressure change.

In order to accurately calibrate the system on the microbar level, we would need to know precisely where the system lies between isothermal and adiabatic. Determining this requires detailed understanding of the dynamics of heat transfer in the system – not only from the air in the chambers to the aluminum walls, but also from one chamber to the other and within each chamber. Even assuming that the system is completely adiabatic, the temperature changes involved are on the order of a *milli*-Kelvin. The danger of attempting to absolutely calibrate the microbarograph system using the incomplete data currently available is that the resulting atmospheric measurements will become less accurate, rather than more so.

Based on this analysis the pressure calibration system was redirected towards relative calibration of the DPTs. Rather than worry about the specific pressure applied, we focused on applying the same pressure differential to all sensors and examining the response relative to a fiducial mean. The pressure readings of the sensors are then adjusted to account for any differences in response as a function of input pressure between the sensors.

μBAR is controlled and data logged using LABVIEW with a USB data acquisition device capable of both analog and digital input/output operations. The DPTs output voltages of 0 to +5 volts in direct proportion to the pressure differential across the sensor diaphragm. These voltages are recorded by the LABVIEW program at a rate of fifty data points per second and recorded along with the dates and times of the acquisition, as well

as the temperature in the DPT enclosure (measured using a thermistor). Calibration data is recorded separately using the same hardware and software.

4.2.2 Weather station

The second principal instrument in the atmospheric arsenal was a Davis Instruments Vantage Pro2 weather station (www.davisnet.com/weather) with the capability to measure temperature, absolute pressure, relative humidity and wind speed and direction among other elements. Of primary concern for this research are the surface wind measurements, because specific wind conditions may be associated with the occurrence of AR or the presence of AGWs. For completeness, temperature and humidity are monitored during observations because temperature fluctuations are known to cause slow distortion of telescope structures and water vapor, as discussed in Chapter 2, plays a small roll in visible wavelength refraction.

Wind speeds and directions are measured using a cup anemometer with a magnetic switch and a wind vane with a potentiometer respectively. Both measurements have a 2.5 second time resolution and directions are measured with a one degree resolution and four degree accuracy while speeds are measured with a one mph resolution and are accurate to ±2 mph. Temperatures are measured using a thermistor every 10 to 12 seconds with 0.1 degree Celsius resolution and 0.5 degree Celsius accuracy. Relative humidity is measured using a thin film capacitor element with a resolution of 1% and an accuracy of 5%. All weather station data are recorded on one minute intervals with means, minima and maxima per interval as applicable and stored in the weather station memory until downloaded via USB to a computer.

173

4.2.3 LIDAR

LIDAR (Light Detection and Ranging) is the optical wavelength analog to radar. ALE, the Astronomical LIDAR for Extinction is located at the UNM Campus Observatory and is designed specifically for the purpose of monitoring atmospheric extinction during photometric astronomical observations. ALE is a range-resolved elastic-backscatter LIDAR with a wavelength of 527 nm that transmits 1500 24 ns long pulses per second into the atmosphere. As the light pulse passes through the atmosphere it is scattered and absorbed by molecules, aerosols, clouds and other atmospheric constituents. Some fraction of each pulse is backscattered to the short and long range receivers (100 mm and 0.67 m telescopes respectively) with the backscattered pulse additionally scattered by the same constituents on the return trip. The fraction of each (carefully measured) transmitted pulse that is backscattered from a set altitude into the receivers as a function of time is closely related to the value of the relative atmospheric extinction.

Below five kilometers ALE is calculated to be sensitive to density fluctuations as small as a 0.1%, where a temperature fluctuation of 0.1 Kelvin at sea level under standard conditions corresponds to a change of approximately 0.03%. The range resolution is a function of the pulse width and is 15 meters for ALE. The actual range is determined from the time-of-flight of the returned pulse and can be determined to well above the troposphere (the region of concern for this research).

At low altitudes (bottom few hundred meters of the atmosphere) the atmosphere is highly contaminated by dust and aerosols such as smoke and other pollutants, especially in the Albuquerque metro area. Although these contaminants may trace some of the more significant boundary-layer structures, the high strength of the backscattered signal from these contaminants will swamp any molecular backscatter, which would normally trace atmospheric dynamics. In general ALEs sensitivity to boundary-layer dynamics is limited by this effect.

As a minor constituent of this research ALE was operated over several nights in conjunction with the initial single telescope operations at the UNMCO. The goals of the LIDAR comparison were to determine if any atmospheric structure observed by ALE was correlated with AR. LIDAR (typically Doppler LIDAR, with vertical wind speed measuring capability) are frequently used by the atmospheric community to observe AGWs throughout the atmosphere.

4.2.4 Atmospheric operations

All science operations of the atmospheric instruments were in tandem with the astrometric observations described in Section 4.1. During the 2007 NOFS observing run, a microbarograph array consisting of two differential sensors and one traditional point source sensor was installed approximately 15 meters to the west of the 1.0 meter telescope (see Figure 49 and Figure 82). The two differential sensors were arranged to measure the pressure differences across orthogonal legs (N-S and E-W) while one port of the point source sensor was open to the atmosphere through one of the legs and the other connected to a needle valve providing a reference volume with a slow leak. The sensor

arms are the white PVC pipes extending from the large box in Figure 82 and extend to the northwest and southeast (across the picture) and to the northeast. Foam block acoustic filters were used during this run (visible in the picture on the ends of the pipes). The weather station was a little over a meter above the surface west of the microbarograph array (on top of the largest white box in the picture). Additional weather data were also obtained from the NOFS weather station mounted outside the 1.55 meter dome.

During the 2008 NOFS observing runs the same location was used for the microbarograph array on all but the final night of May observations (the array was moved to the parking lot north of the 1.0 meter due to wet ground). Two differential legs were used during the April run and three differential legs plus a point source sensor during the May run. The array was laid out with inlets at three corners of a roughly equilateral triangle such that differences were measured across the sides of the triangle. In April one sensor measured a north-south pressure difference while the second measured a northeast-southwest pressure difference (see Figure 83). In May differences across all three sides of the triangle were measured while the point source opened to the atmosphere along the north sensor arm. The expansion chamber acoustic filters were used during all 2008 runs and make up the entire sensor arms visible in the picture.

The weather station was mounted approximately four meters above the surface on a pole south of the microbarograph and 1.0 meter dome (behind the scaffolding visible in the background of Figure 83) during the 2008 observations.

Figure 82. μBAR setup during the 2007 NOFS observing run. The columns to the left of the image run N-S. The 1.0 meter dome is off the image to the left.

Figure 83. April 2008 μBAR setup with expansion chamber acoustic filters. The 1.0 meter telescope dome is to the left of the image.

177

Microbarograph use was discontinued during the UNMCO observations due to the general lack of consensus between the µBAR measurements and observed AR during the NOFS runs; however, a weather station mounted above the east wall of the observatory was operated continuously throughout all UNMCO operations. During the initial operations of the east 10-inch telescope at the UNMCO ALE was being operated as part of an unrelated photometry experiment and the resulting data from several of these nights compared to the simultaneous astrometric data. ALE is located several meters east of the two 10-inch telescopes in the observatory courtyard (see Figure 84) and was pointed at the zenith during all operations.

Figure 84. UNMCO courtyard and three of the four courtyard instruments. The dome of the 10"E telescope used in this research is in the foreground and ALE is housed in the furthest dome. A Spectrophotometer (not used in this investigation) is housed in the middle dome.

178

4.2.5 Atmospheric data processing

The largest factor in the atmospheric data reduction is the microbarograph data. Data from the weather station is retrieved either directly from the station in ASCII (tab table) format or acquired during the course of pressure operations through a LABVIEW program and written to a file. Lidar data are obtained in a processed state from other scientists in the research group, yielding backscatter as a function of time and altitude.

μBAR data processing (see Appendix H, code *mbmaster.m*) begins with data from the calibration run, if one was taken. The calibration data include time and pressure data during operation of the pressure calibration chamber. Pressure data are recorded as voltages from 0 to +5 volts and converted to pressures based on the full scale setting of the DPT using $p(mb) = FS(volts - 2.5)$, where FS is the selected full scale setting (the three user selectable ranges are $FS = \pm 0.1245$ mb, $FS = \pm 0.249$ mb and $FS = \pm 0.498$ mb). The mean pressure reading from the sensors used at each data point is determined and a linear fit of individual sensor pressure to the mean pressure is calculated. This fit will be removed from the science pressure data to correct for any differences between the sensors.

Temperature data are recorded throughout both the calibration sequence and the science measurements. Both the calibration pressure readings and the science pressure readings are corrected for temperature errors based on the empirically determined temperature errors for each sensor. Science pressure data are converted to millibars in the same manner as the calibration pressure data. Data acquired by the weather station are cropped

to the same time range as the science pressure data. The final state of the reduced data is vectors of times and pressures from the microbarographs, temperatures (and times recorded) in the microbarograph electrical box, and vectors of weather station recorded times, wind data, temperatures and relative humidity.

With this unique conglomeration of astronomical and atmospheric instrumentation and observations we accomplished a groundbreaking study of anomalous refraction that has led to new insights into the cause and nature of AR, error sources in astrometric measurements, and the nature of the nocturnal atmosphere.

5. Results

When this research began, the prevailing opinion of the astrometry community held that AR was caused by kilometer scale coherent structures in the atmosphere, *e.g.* AGWs. Thus the experimental phase of the project was entered with the expectation that large-scale atmospheric gravity waves would cause multiple telescopes at the same observatory to see the same anomalous refraction. Although the anomalous refraction observed was highly consistent in character with all previous observations and also highly consistent across telescopes, the expected correlation between simultaneous observations has not been evident, all but ruling out atmospheric gravity waves. Rather than delving into the precise quantitative relationship between AGWs and AR, the latter parts of this research have been devoted to elucidating the true source of AR and its characteristics.

5.1 Characteristics of AR

The astrometric data were obtained over the course of three years and more than two dozen nights of observation on a wide array of telescopes. This dataset includes about ten nights of single telescope data, 15 nights of two telescope data and one night of data on three telescopes. For ten of those nights observations were made at the US Naval Observatory in Flagstaff, Arizona (NOFS) on meter-class telescopes. The rest of the data were obtained on off-the-shelf 10 inch telescopes at the UNM Campus Observatory in Albuquerque, New Mexico (UNMCO). The defining quality of all of these observations is the ubiquitous occurrence of anomalous refraction. Anomalous refraction is a continuously occurring phenomenon that has been observed in every instance where a telescope is observing in a mode sensitive to its effects, regardless of telescope design or observatory location.

In summary, the anomalous refraction observed during the course of this research can be described as loosely periodic astrometric positional fluctuations with two broad fluctuation timescales. Long period motions, which may be partially caused by telescope motion (mechanical or thermal), have periods of several tens of minutes to hours and amplitudes ranging from several tenths of an arcsecond to as large as two arcseconds. Short period motions have periods ranging from one minute to roughly 20 minutes and amplitudes of generally less than half an arcsecond.

A typical example of the residuals characteristic of anomalous refraction is shown in Figure 85. Both minute timescale and tens of minute timescale motions occur throughout the two and a half hour observation made using the UNMCO 10"E telescope and are further illustrated by Figure 86. Another characteristic of anomalous refraction is the disparity between the residuals in right ascension and the residuals in declination. In general the RA and Dec residuals are only rarely correlated.
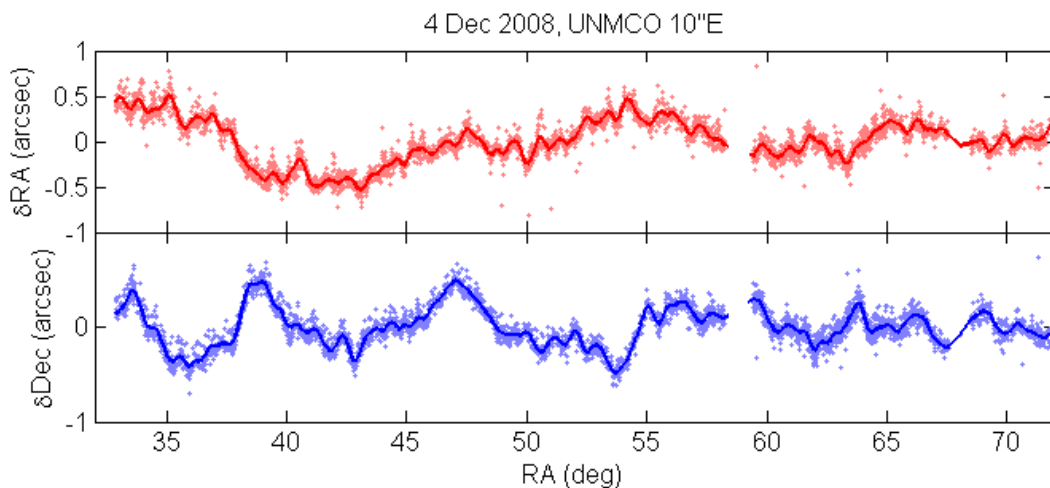

Figure 85. Example of characteristic anomalous refraction residuals. Residuals are differences between RA and Dec image positions and associated catalog positions for each star. Solid line is a smoothing spline fit to the residuals to highlight the coherent motions.
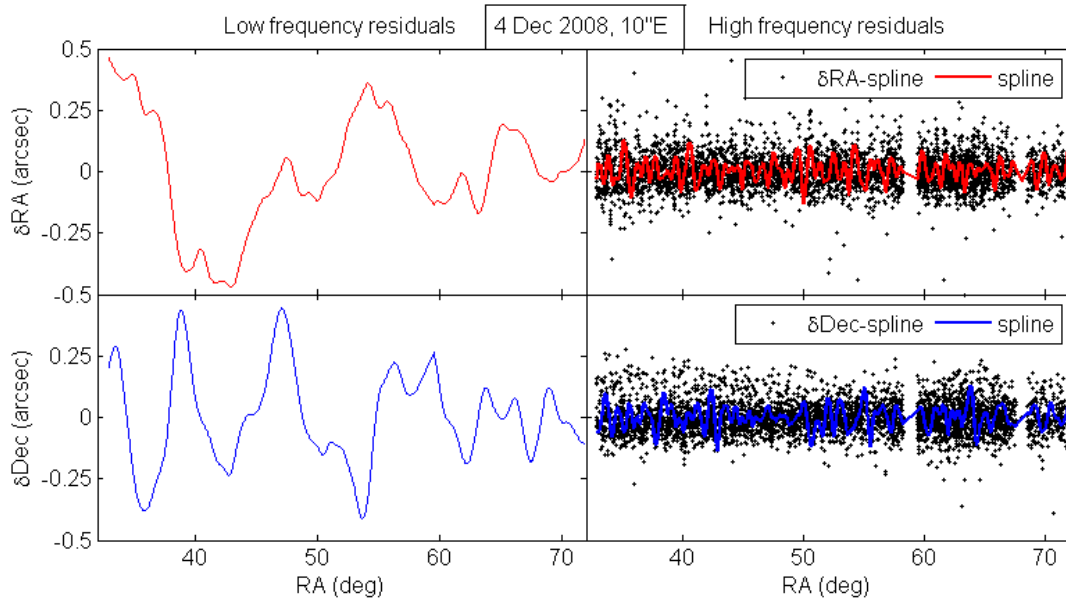
182

Figure 86. Smoothing spline fits to the low and high frequency components of the residuals in Figure 85 (taken with the UNMCO 10"E telescope). The left two plots are spline fits to the RA and Dec residuals with periods of tens of minutes. The right hand plots are spline fits to the residuals with periods of minutes, the splines in the left plots having been subtracted. For comparison, the noise in the residuals (residuals with all AR removed) is shown as the black points in the right-hand plots.

The timescales of anomalous refraction are further illustrated through a study of the frequency spectra of the residuals. Figure 87 shows the frequency spectra in both RA and Dec from the same data taken at the UNMCO on December 4th 2008. Dominant periods of ten to twenty minutes have amplitudes an order of magnitude larger than the dominant periods in the few minute range (enlarged in the inset). For this particular example the image integration time is a little more than 100 seconds. Periods of image motion of order the integration time or shorter are damped as evidenced by the reduced power in the spectra around 0.01 Hz and higher frequencies. A linear fit to the power spectra in log-space in the period range 100 seconds to 3600 seconds (examples from the December 4th UNMCO 10" data and the May 10th 2008 NOFS 1.0 meter data are given in Figure 88

and Figure 89 respectively) indicates a power law dependence variable both between RA and Dec residuals and from night to night with typical values of $F(f) \propto f^{-2/3} - f^{-2}$. The power laws are consistent to within this range of variation regardless of the telescopes or observatories used. The power law dependences obtained for most of the NOFS and UNMCO nights are plotted in Figure 90.
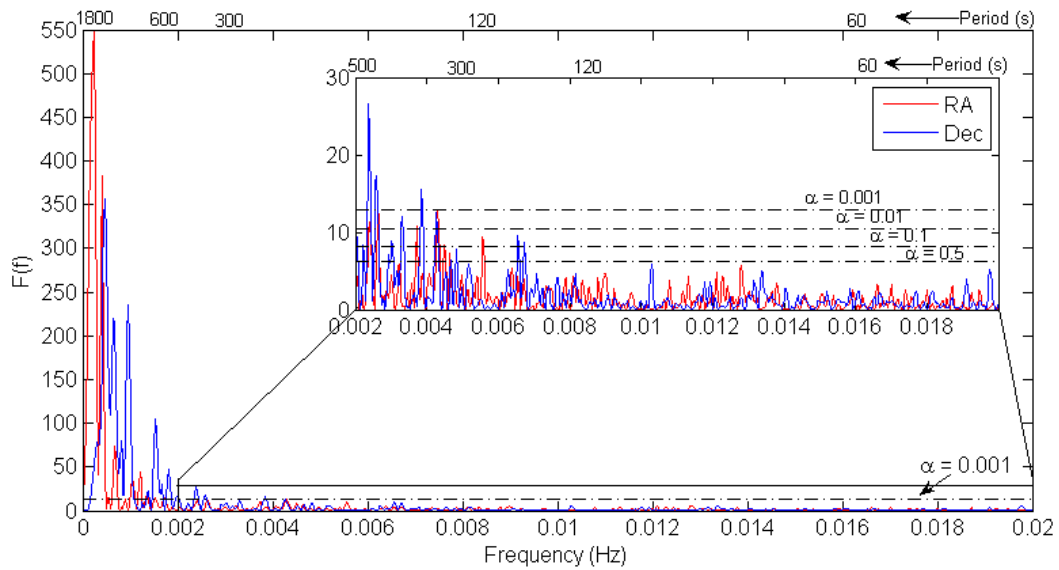
Figure 87. Periodograms of RA and Dec residuals from 4 Dec 2008. Significance levels are indicated by the dashed lines with values indicated by α. The significance is the probability a given power level of a frequency will be produced in a random signal. Values greater than α=0.05 are considered significant.
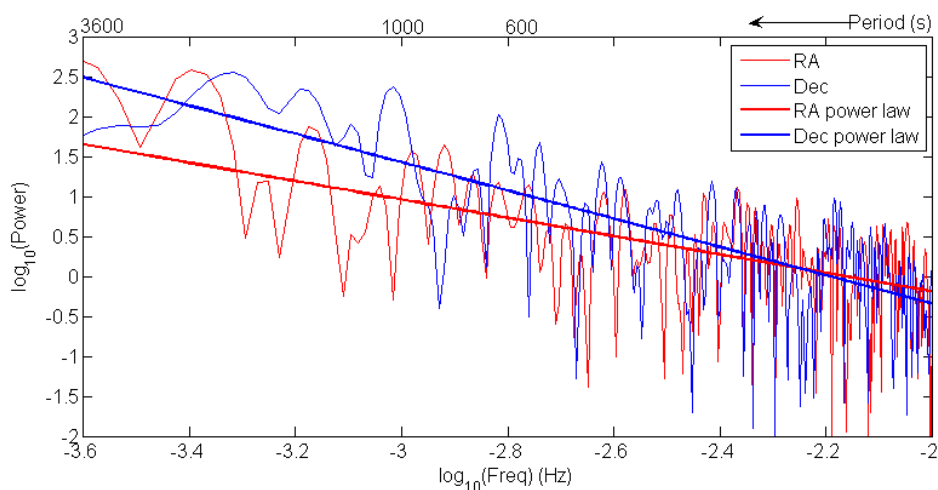
Figure 88. Log-space plot of the power spectra from 4 Dec 2008. Linear fits to log(power) vs. log(freq) for both RA and Dec are indicated by the red and blue lines respectively.

184
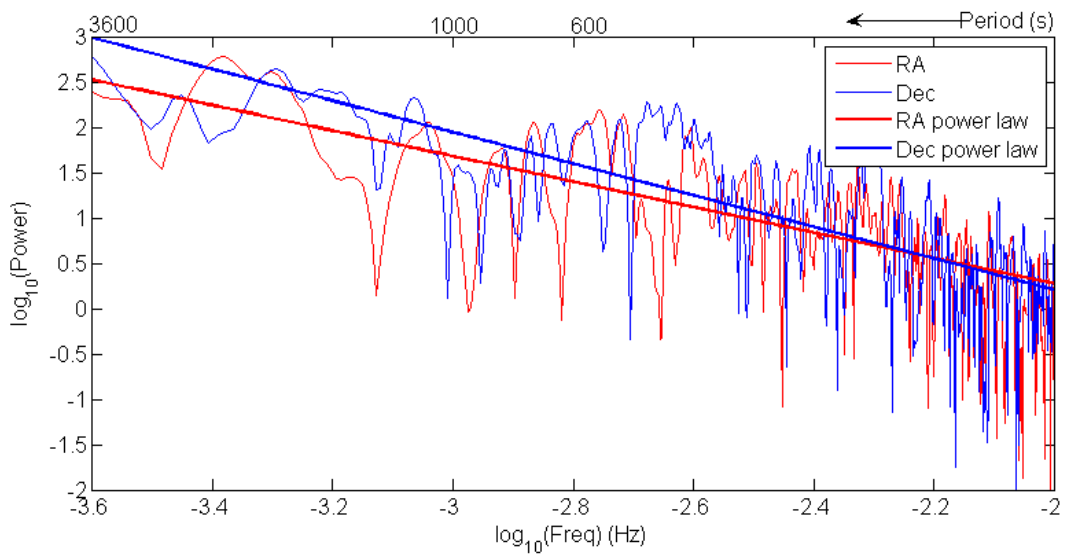
Figure 89. Same as **Figure 88** for observations made using the NOFS 1.0 meter on 10 May 2008.
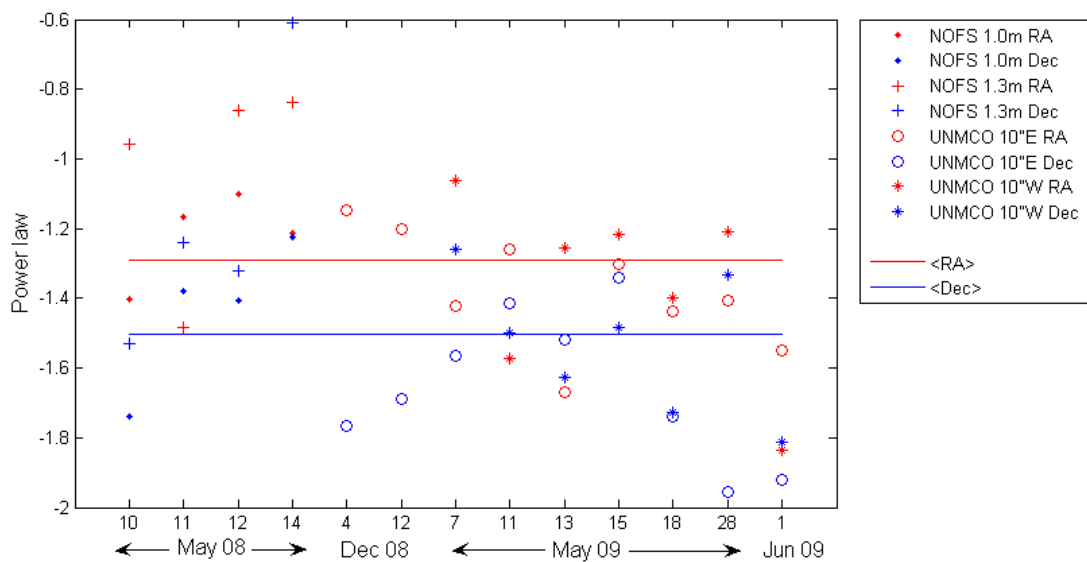
Figure 90. Power law dependencies for most NOFS and UNMCO observations. Solid lines indicate the average power law dependence for all included RA residuals (red) and the average power law dependence of all Dec residuals (blue).

5.2 Consistency Among Observations

The feature of the astrometric data that must be noted before any further analysis is the consistency between the anomalous refraction observed both at NOFS and the UNMCO

185

as part of this research and that which has been observed by other telescopes (Figure 91). All observed anomalous refraction consistently matches the description above regardless of when it was observed, the observatory location or nature of telescope with which it was observed. Figure 90 also indicates that the power law dependence of the frequencies of AR is consistent between telescopes and observatories to within the night to night variability. Additionally, over the course of two dozen plus nights of observation on varying telescopes at both NOFS and the UNMCO as well as analysis of several dozen datasets from the Sloan telescope, there have been no nights or even fractions of nights where anomalous refraction is *not* present.
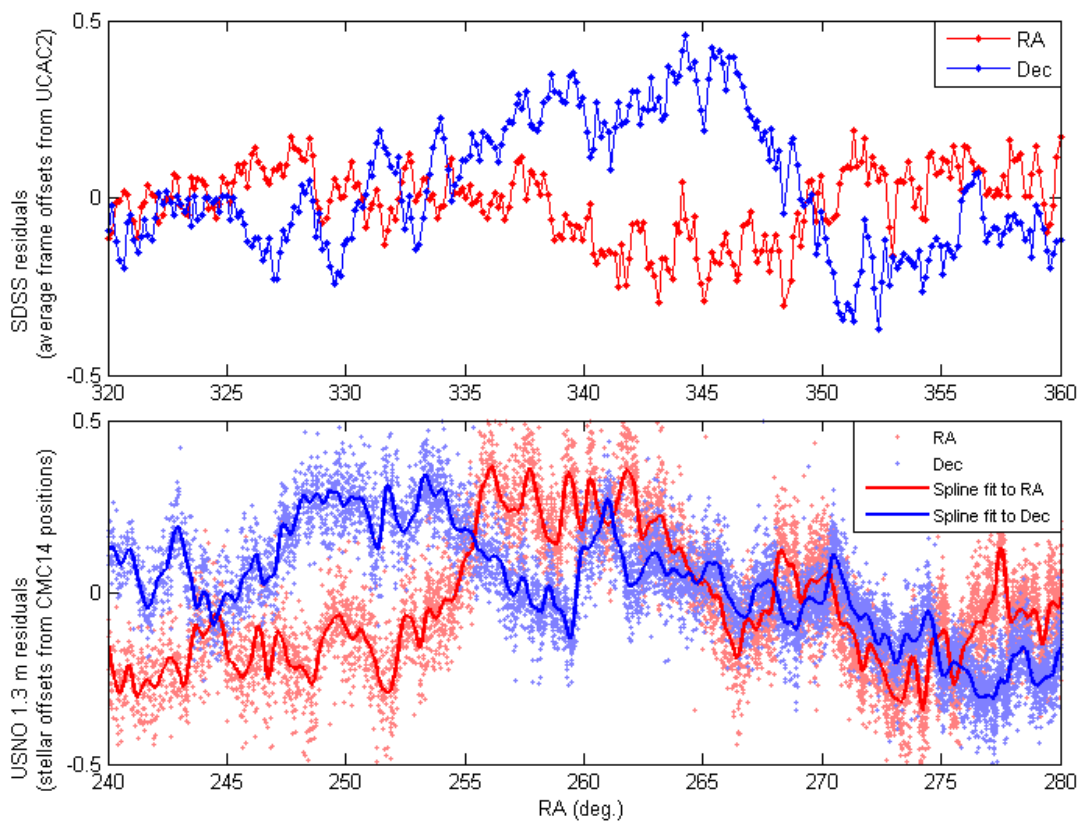
Figure 91. Example of NOFS residuals obtained in 2008 and SDSS residuals from several years earlier showing the characteristic consistency between anomalous refraction observed by different telescopes at different sites. Fifteen degrees of RA corresponds to one hour.

5.3 Comparison with Surface Weather

The continuous occurrence of AR nullified the question of whether AR's occurrence or non-occurrence could be linked to surface weather conditions. Comparisons between the characteristics of AR and surface weather conditions consistently indicate no direct relationship between the amplitude and frequencies of AR and wind speed, wind direction or temperature on all nights of observation (see Figure 92 and Figure 93). Winds during the NOFS observing runs ranged from calm to 20 mph and were primarily from the northeast and southwest (Figure 94). Similar conditions persisted during the UNMCO observations.
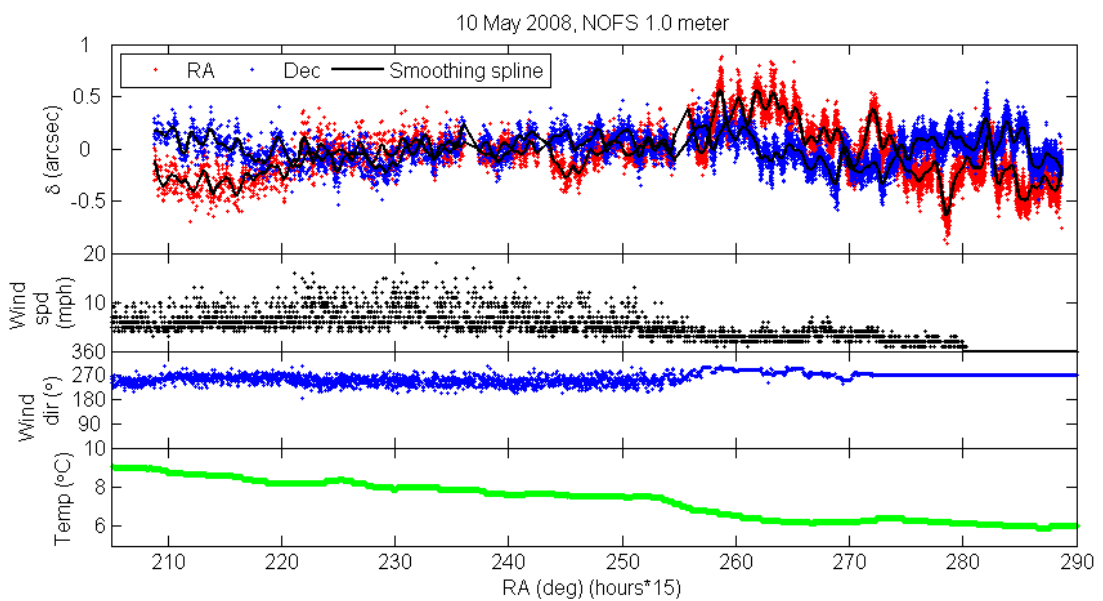
Figure 92. NOFS 1.0 meter residuals for May 10th, 2008 and concurrent surface weather conditions. The amplitudes of the residuals deceptively appear to increase around an RA of 255 degrees due to a combination of the increased stellar density and offset of the RA and Dec residuals.

187
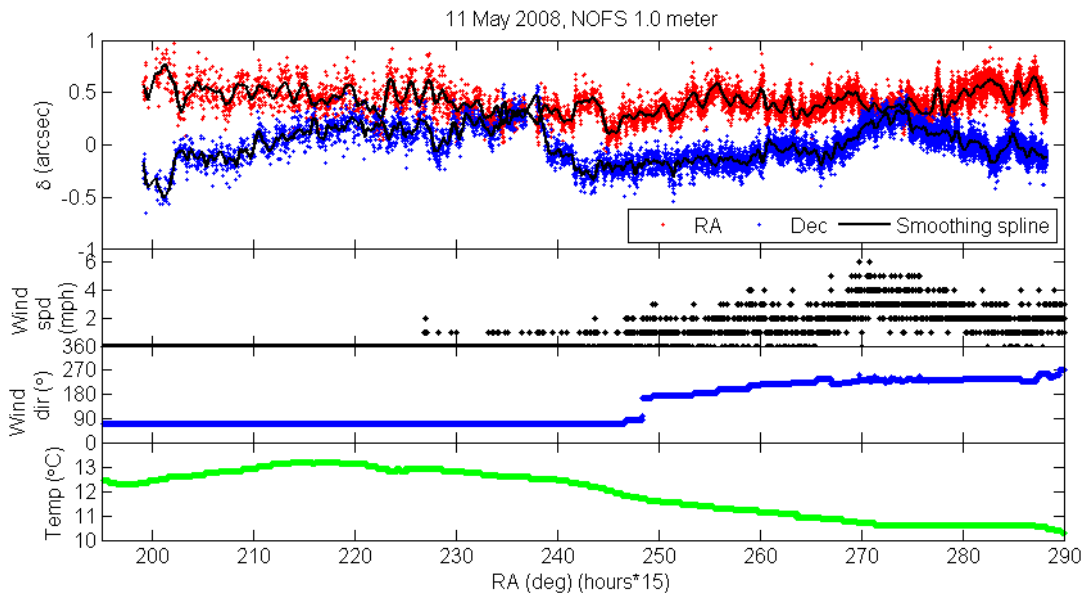
Figure 93. NOFS 1.0 meter residuals for May 11th, 2008 and concurrent surface weather conditions.
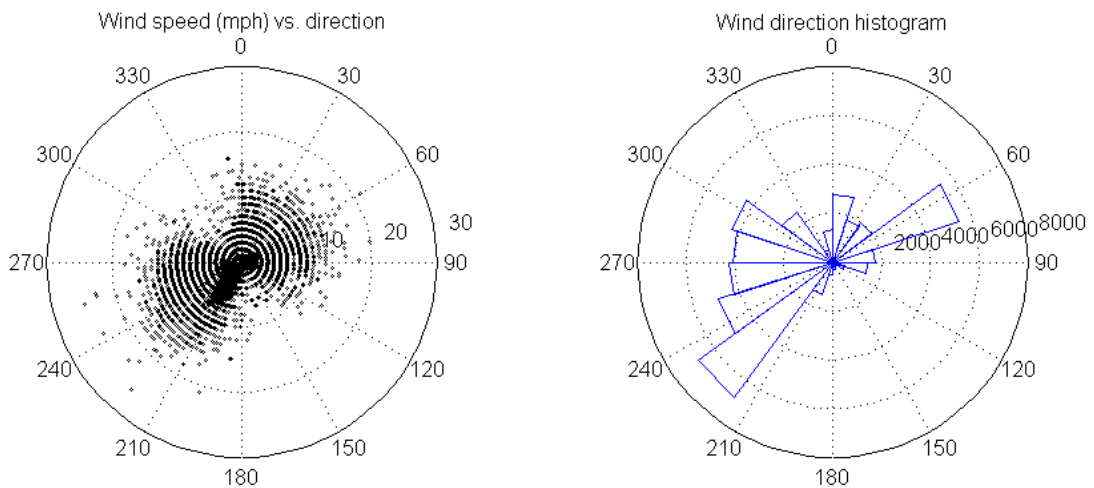
Figure 94. Wind conditions at NOFS during the April and May 2008 observations.

5.4 Comparison with μBAR

Comparisons with pressure data obtained using the microbarograph array during Flagstaff operations also revealed no correlations between concurrent observations. Frequency characteristics of the pressure traces are similar to those of the residuals, with the same power law dependences in the exposure time to one hour period range (Figure 95). The

188

actual pressure and residual traces, however, show no relationship between the amplitudes, phases and timescales of pressure changes in any of the microbarograph legs and the same features of the residuals in either telescope at any given time (Figure 96 and Figure 97). Of particular note in Figure 96 is the region between 270 and 290 degrees where the nature of the microbarograph pressure fluctuations changes dramatically without any change in the characteristics of the anomalous refraction.
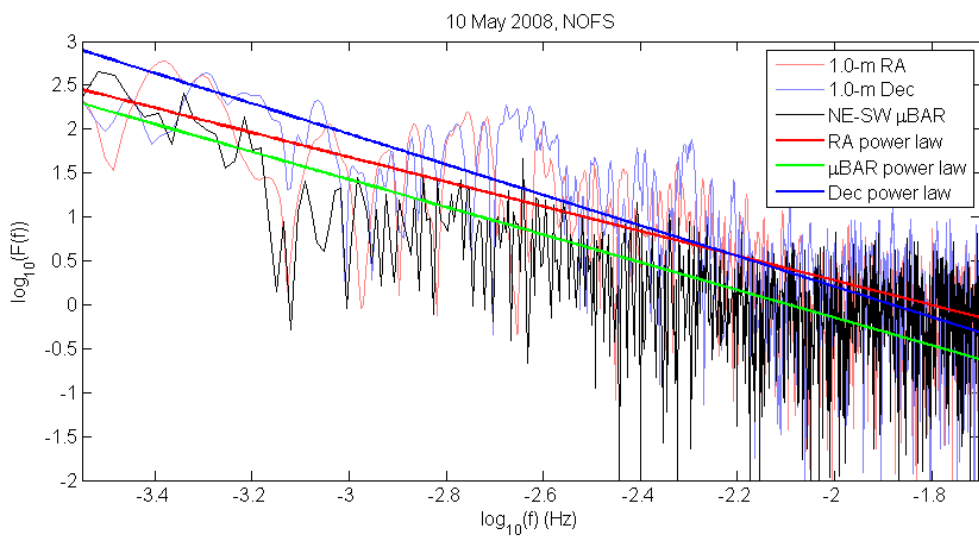
Figure 95. Power spectra of RA and Dec residuals from the NOFS 1.0 meter telescope on 10 May and power spectra of one differential microbarograph during concurrent operations.
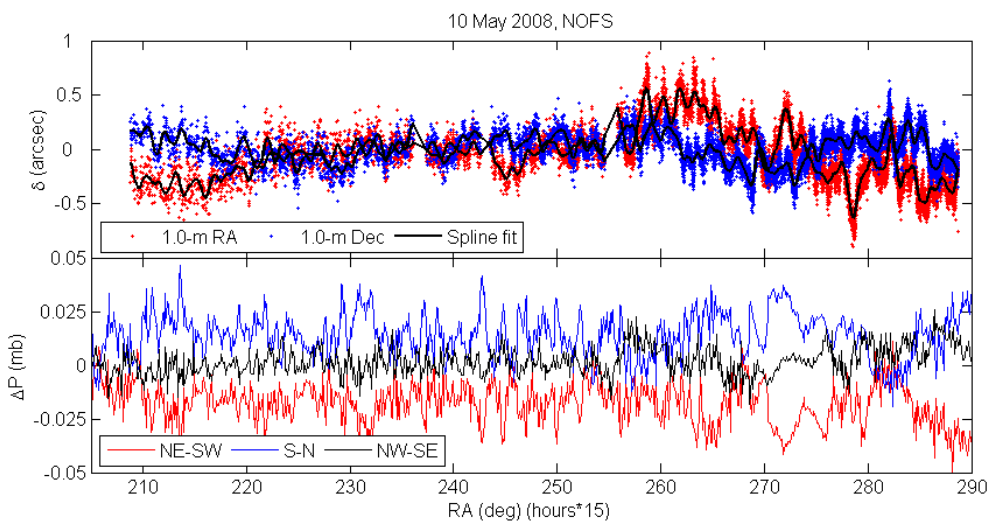
Figure 96. NOFS 1.0 meter residuals for the night of May 10, 2008 and concurrent microbarograph pressure traces from the three differential microbarographs.

189
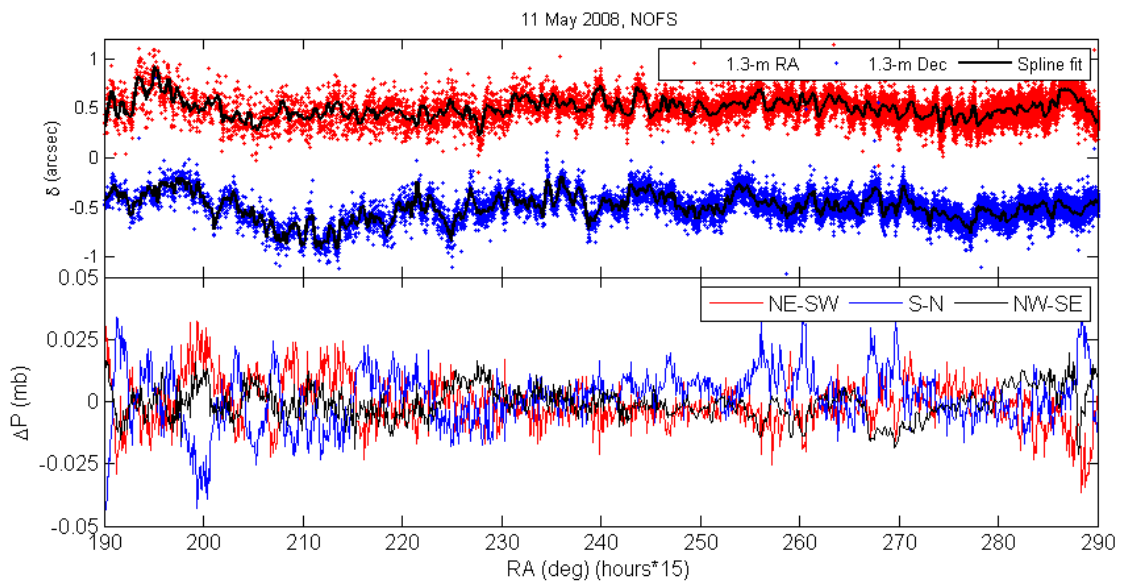
Figure 97. NOFS 1.3 meter residuals for the night of May 11, 2008 and concurrent microbarograph pressure traces from the three differential microbarographs.

5.5 Comparison with Cloud Cover

The only observed weather feature that was found to have any effect on the residuals was the presence of transparent cloud cover. A bank of thin stratus clouds remained over the Naval Observatory for over an hour and a half during one night of April observations and another night was cut short by thickening cloud cover. In both cases, the noise in the residuals and the amplitudes of the anomalous refraction increased when clouds were present (Figure 98). The increase in noise was due to reduced signal-to-noise ratios while the increased residual motions may be explained by non-flat backgrounds biasing the centroids obtained by Source Extractor.
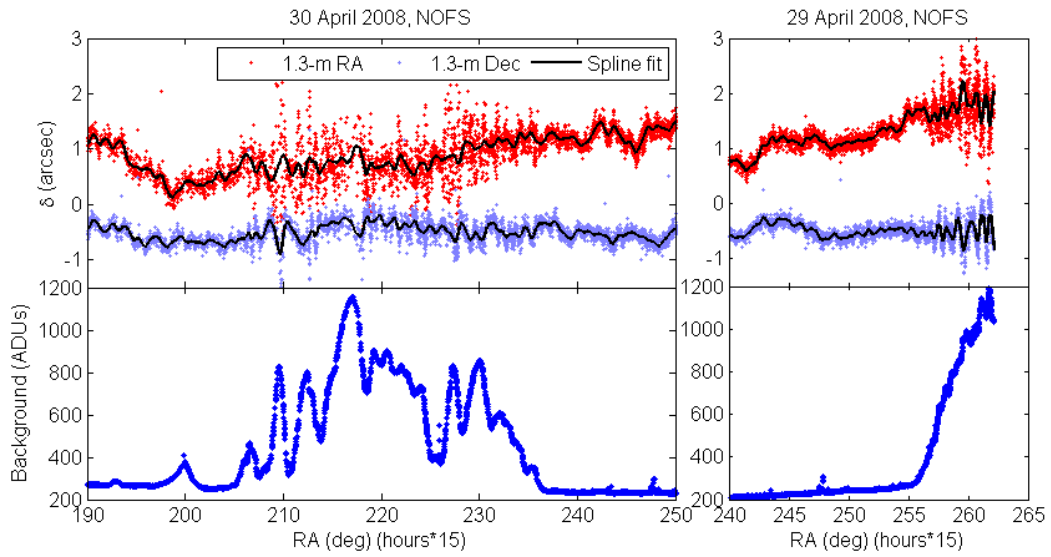
Figure 98. Effect of cloud cover (indicated by increased image background levels) on residuals. Clouds on the 30th of April were thin stratus that significantly reduced transparency, although sufficient star density remained to allow continued observing. Clouds on the 29th of April were a bank of thick stratus that moved in early in the morning and quickly reduced transparency causing an early end to observing.

5.6 NOFS Multiple Telescope Observations

The key results obtained during the observations involve the comparisons between residuals from simultaneous observations on multiple co-located telescopes. Four nights of observations on the NOFS 1.0 meter and 1.3 meter telescopes, with a separation of approximately 300 meters resulted in residuals with very few clear correlations in either RA or Dec. The residuals from one night of two-telescope operations at NOFS are shown in Figure 99 and appear generally unrelated on both short and long timescales.

For a more rigorous determination of the level of coincidence between residuals, the separation between the data as a function of time is calculated in terms of the precision with which the residuals are known at each time (σ). Prior to this calculation, deviations in the residuals with periods of order an hour or more are removed by applying a high-

191

pass filter (in addition to the initial cubic fit that is removed from most datasets) that has the same frequency filtering characteristics for all datasets (Figure 100). A low-pass filter (a square-hat filter with its sharp edges replaced by a half-cosine fall-off on either end) applied to each dataset leaves only the hour or longer periods. The resulting smoothed dataset is subtracted from the unfiltered data to remove the hour or longer periods, leaving only the residuals with minute to tens of minutes periods.
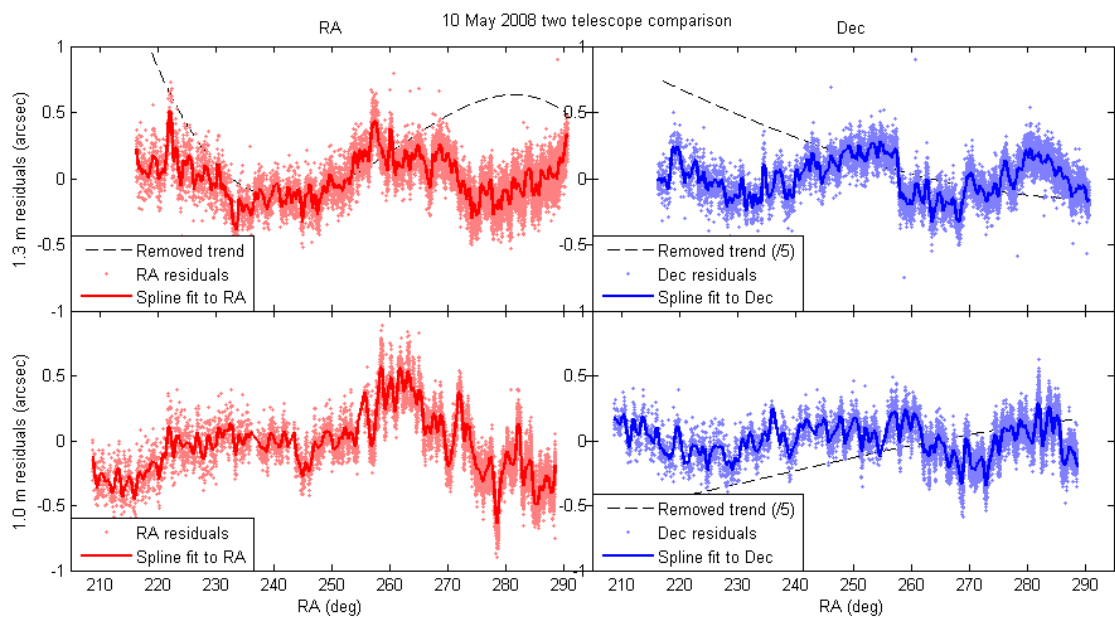
Figure 99. Comparison between 1.3 meter (top) and 1.0 meter (bottom) NOFS residuals. Dotted lines indicated large scale trends that were subtracted from the residuals.

For NOFS and UNMCO data, where each residual point corresponds to the positional offset between a specific star and its catalog counterpart, each run is divided up into bins containing roughly 20-100 stars and sized such that residual motions within a bin are primarily linear (typically $0°.25$ bins are used, see Figure 101). For SDSS data, where each residual point corresponds to an average frame offset, individual points are used instead of bins. The level (in σ's) of coincidence (correlation) in a bin, α_c, is defined as

the difference between the means of the residuals of the two datasets in the bin divided by

the precision to which those means are known in that bin, *i.e.*

$$\alpha_c \equiv \frac{\langle\delta_1\rangle_{bin} - \langle\delta_2\rangle_{bin}}{\sqrt{\left(\frac{\sigma_1}{N_1}\right)^2_{bin} + \left(\frac{\sigma_2}{N_2}\right)^2_{bin}}},$$

where δ_i denote the residuals of each dataset, σ_i are the standard deviations of each set

and N_i are the number of points. The residuals in a given bin are highly correlated if

$\boldsymbol{\alpha_c < 2\sigma}$. Figure 102 shows an example of the application of this metric and general

agreement with qualitative assessments of the level of correlation of the data.
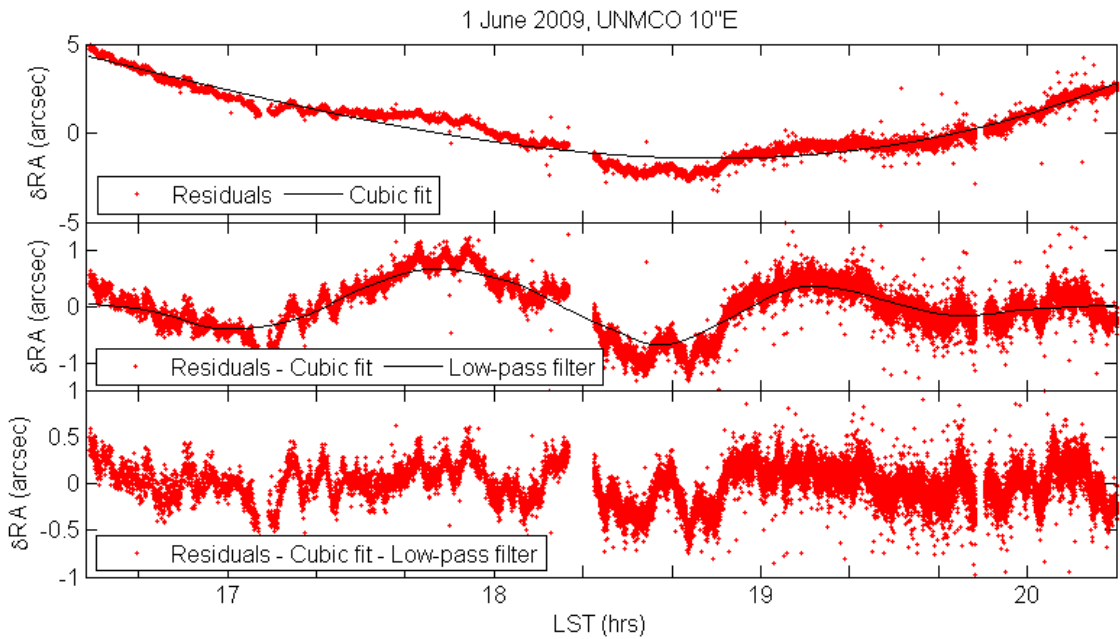
Figure 100. Filtering applied to all datasets prior to quantitative calculation of the level of correlation. A cubic fit removes slow (typically diurnal) motions. Remaining motions with timescales of hours are removed by first filtering out the higher frequencies, then subtracting the filtered data from the original data, leaving only the higher frequencies.
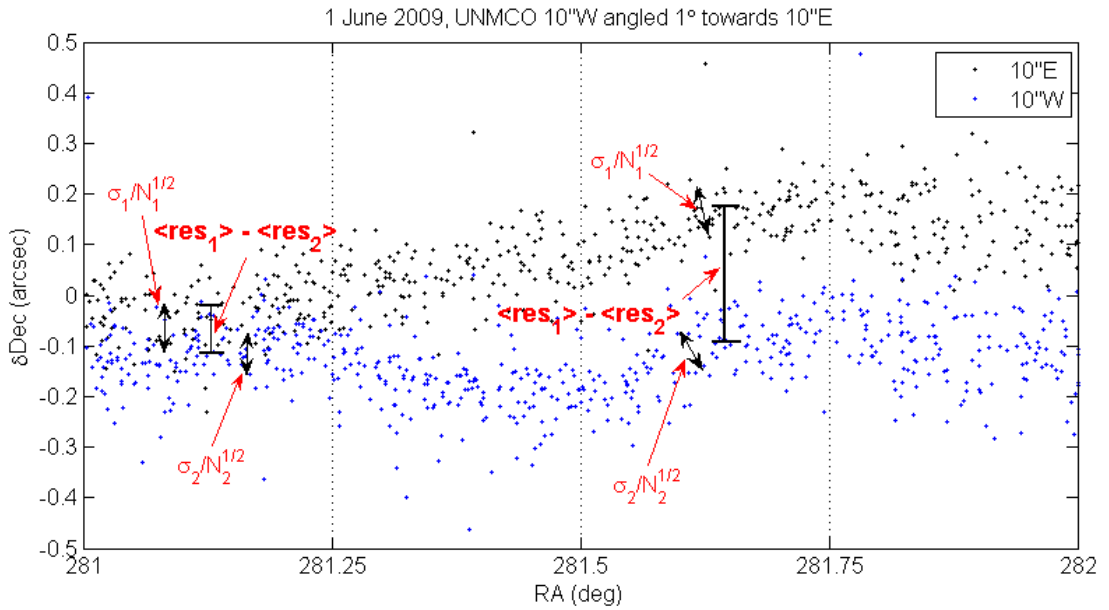
Figure 101. Illustration of the quantities used in the calculation of residual coincidence. In each quarter degree bin the difference in mean residual positions is divided by the precision to which those means are determined, which is a function of the standard deviation of the residuals in the bin divided by the square root of the number of points in the bin.

The level of coincidence for a full night is defined by the fraction of bins coincident within two sigmas. Offsetting one set of residuals by some RA and comparing the resulting coincidence fraction with that of the simultaneous residuals provides an internal reference of the coincidence fraction for an unrelated set of residuals. This is applied in our analysis by calculating the coincidence fraction for a range in RA offsets from -10° to +10°. If the two sets of residuals are correlated, a plot of the fractional coincidence as a function of offset angle (Figure 103) will show a significant peak at zero offset and a sharp fall-off in fractional coincidence with increasing offset, bottoming out at a nominal value indicative of chance coincidences in uncorrelated data. If there is a phase lag either between CCDs (on SDSS) or telescopes, the peak in the fractional coincidence verses

offset plot will occur at a non-zero offset and will indicate the time-lag between when each detector sees a particular coherent disturbance (as in a cross-correlation).
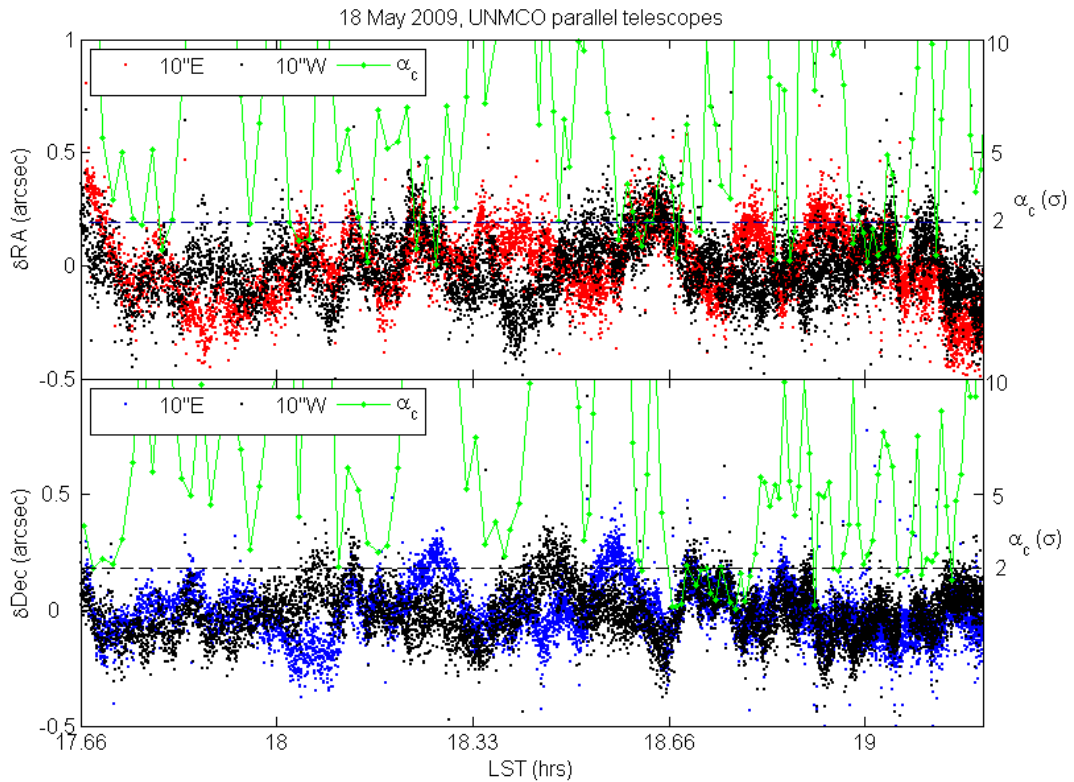
Figure 102. Simultaneous residuals from co-located telescopes and the level of correspondence, α_c. Values of α_c less than two sigmas indicate significantly corresponding residuals.

Applying the correlation metric to the SDSS datasets provides a clear example of how we should expect the quantitative measure of correlated data to appear (Figure 103). Examining the fractional correlation as a function of distance between CCDs on the SDSS focal plane further defines the level of coherence as a function of spatial scale (Figure 104). The correlation of the SDSS residuals as a function of CCD separation was discussed at length in Chapter 3 with the general conclusion that adjacent CCDs had highly correlated residuals while opposite CCDs, while still correlated, showed clear differences between the residuals. The dependence of the fractional correlation on CCD

195

separation is highly consistent across different observations, varying in amplitude, but maintaining the same shape. From Figure 104 it is now clear that highly correlated residuals will have fractional coincidences of 60% - 80% while fractional coincidences of 45% or more indicate residuals that are still correlated, but to a lesser degree. Rather than absolutely pinning numbers to the degree of correlation, however, the best measure of correlation is whether the fractional coincidence changes significantly when simultaneous datasets are offset.
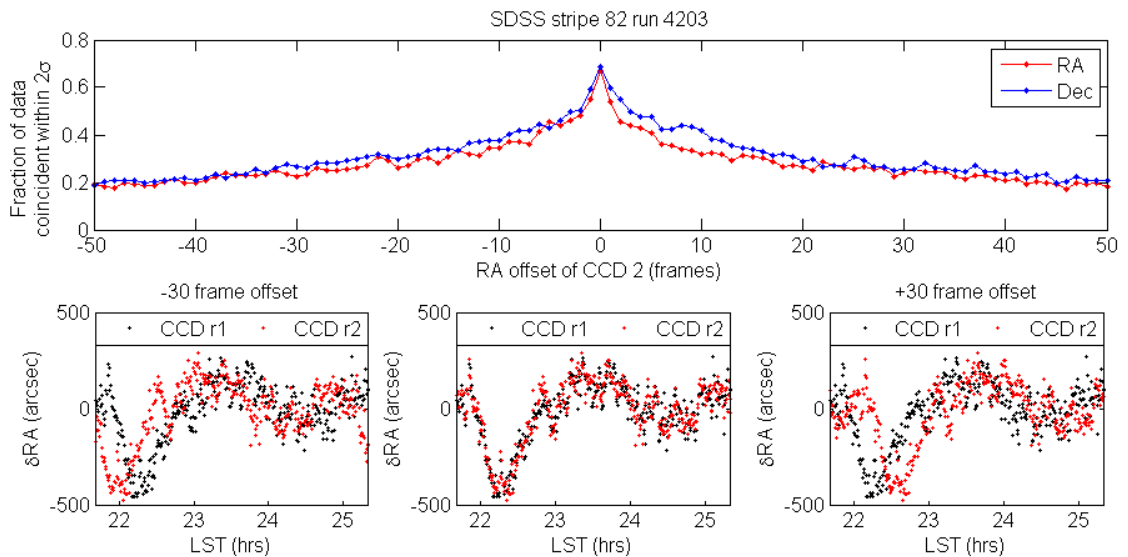
Figure 103. Fractional coincidence as a function of offset (in frames, with each frame 0°.15 RA long) for adjacent CCDs in the Sloan focal plane. The high correlation of the two datasets is evident in both the large fraction of frames coincident to within 2σ at zero offset and the large fall-off in the coincidence fraction with offset. Examples of the data used with zero offset and offset by ±30 frames (±4°.5) are shown in the bottom three panels.

Applying the correlation metric to NOFS data confirms the previous statement that simultaneous residuals obtained on the 1.0-m and 1.3-m telescopes are completely uncorrelated (Figure 105 and Figure 106). The apparent peak in coincidence in declination in Figure 105 loses significance in light of the lack of any clear fall off as in the SDSS data, the low prominence and the lack of any corresponding peak in RA.

196
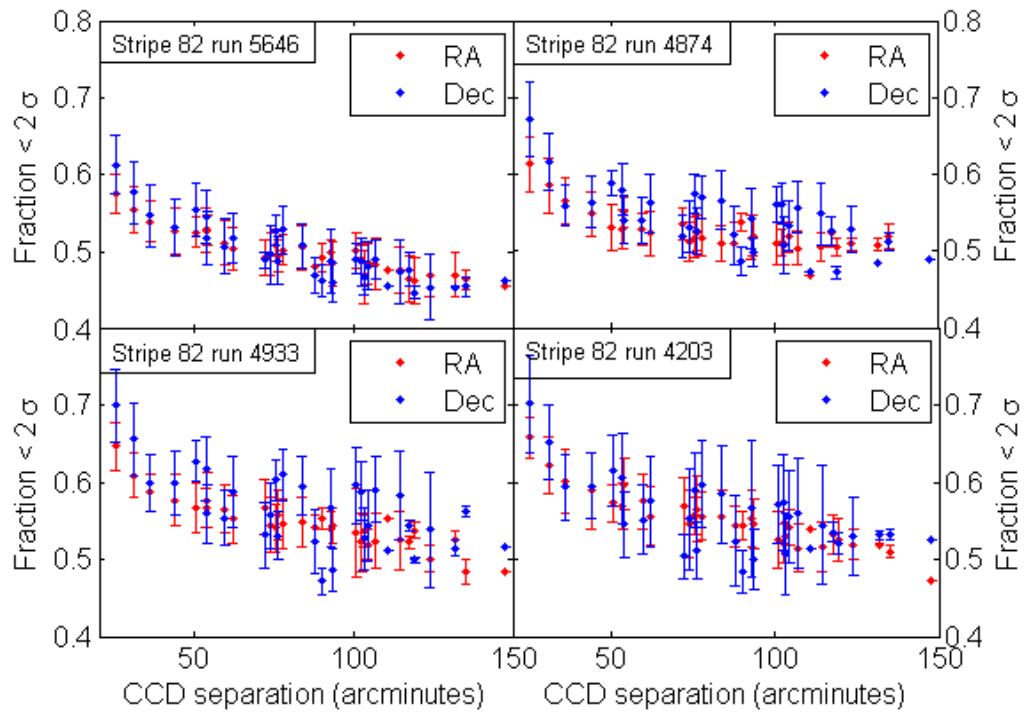
Figure 104. Fractional coincidence of SDSS residuals as a function of the distance between CCDs on the focal plane for four equatorial stripe runs. Each point is the mean fractional coincidence for all unique combinations of CCDs with a given separation (*i.e.* separation between the first and second CCDs in row one is the same as that between the second and third and so forth), while the error bars indicate the standard deviations.
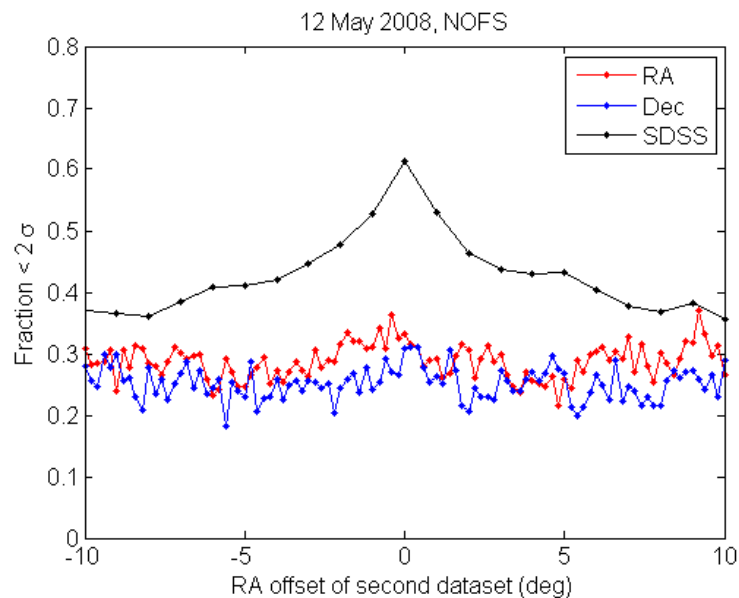
Figure 105. Fractional correspondence of residuals from the NOFS 1.3-m and 1.0-m telescopes on May 12, 2008 and an example of SDSS two CCD correspondence for comparison.
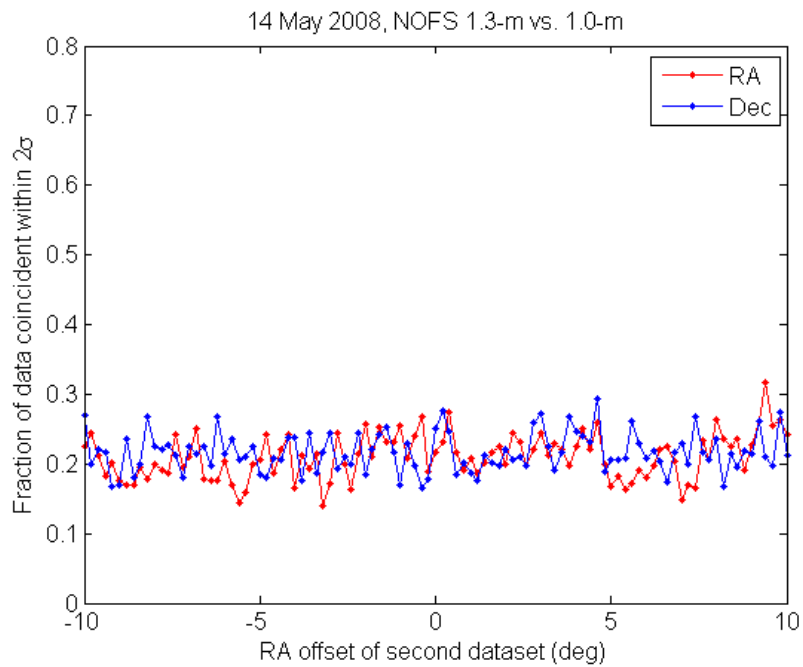
Figure 106. Fractional correspondence of residuals from the NOFS 1.3-m and 1.0-m telescopes for May 14, 2008.

One night of simultaneous observation on the two telescopes above as well as the 1.55 meter, separated by 50 meters from the 1.3 meter, also showed no significant correlations between the three telescopes (see Figure 107). Quantifying this result once again confirms that no correlation exists between any of the three telescopes, even the telescopes with the closest spacing (Figure 106, Figure 108 and Figure 109).

Comparing the frequency trends of the two primary NOFS telescopes (1.0-m and 1.3-m) on any of the nights with concurrent observations gives power law dependencies that are similar, but variation between the telescopes is generally of the same magnitude as night to night and RA to Dec variation (see Figure 90 and Figure 110). Notable differences in power at many frequencies in the few minute to tens of minutes period range are clear in Figure 110.
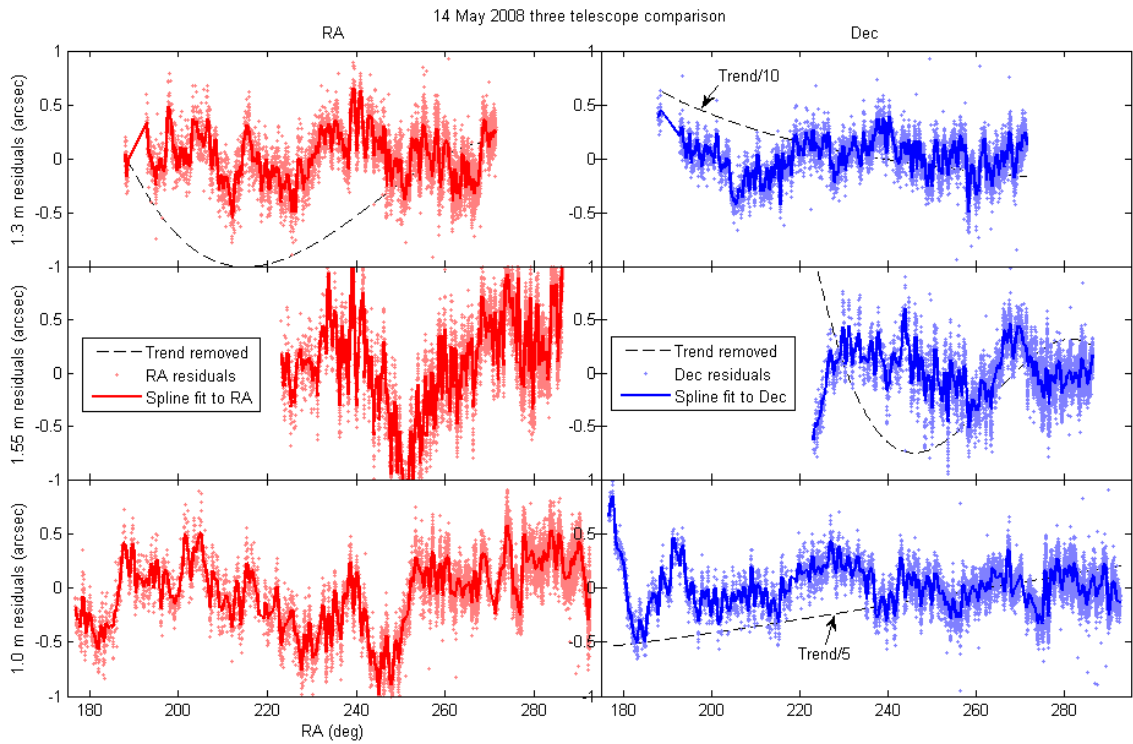
Figure 107. Comparison between all three NOFS telescopes. Dashed lines (scaled as indicated to fit in the figure) indicate large scale trends removed from the residuals.
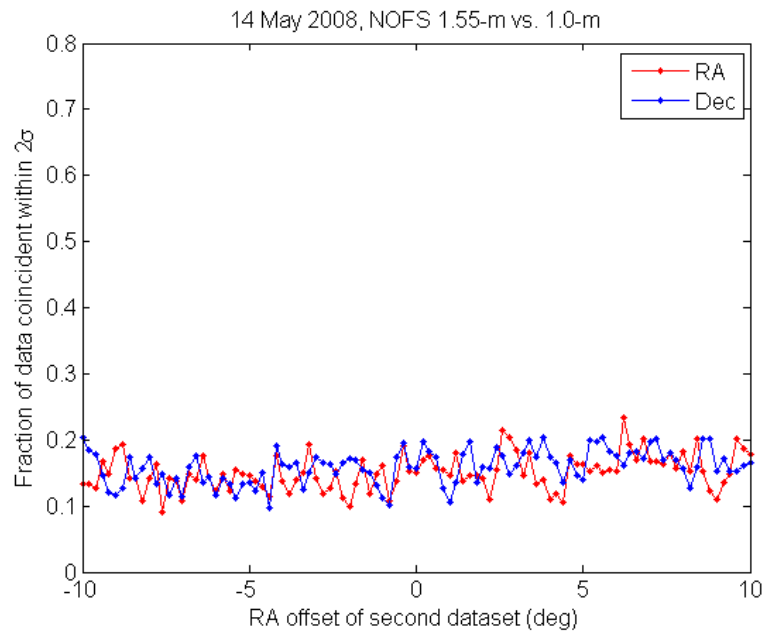
Figure 108. Fractional coincidence between simultaneous residuals from the 1.55-m and 1.0-m telescopes on May 14, 2008.
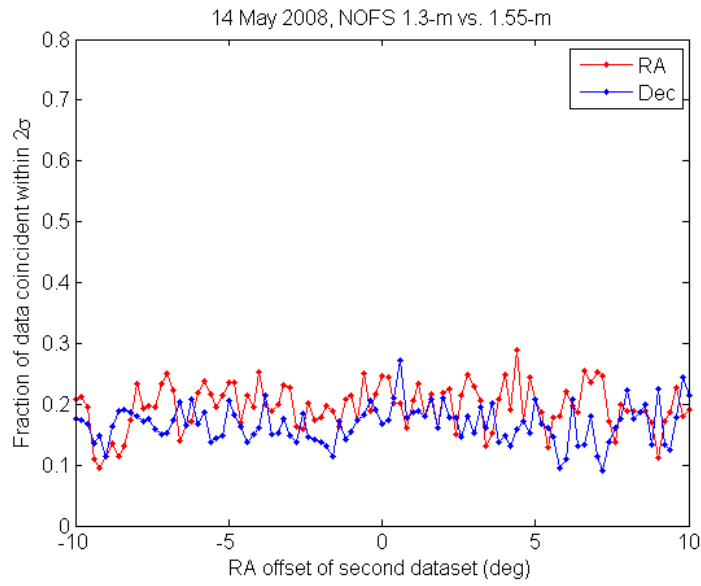
Figure 109. Fractional coincidence between simultaneous residuals from the 1.55-m and 1.3-m NOFS telescopes on May 14, 2008.
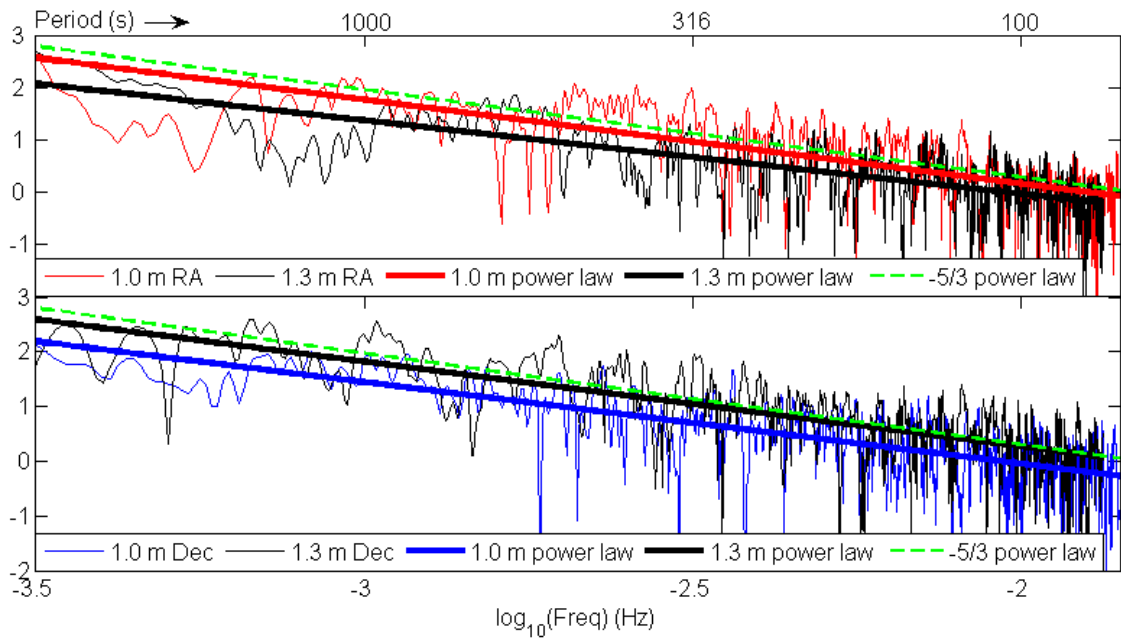
Figure 110. Power law dependence comparison for the NOFS 1.0 m and 1.3 m telescopes on 11 May 2008. Top: RA comparison with a -5/3 power law indicated by the dashed green line. Bottom: same as top for Dec. Power law dependencies are determined by a linear fit to the log-space frequency spectrum in the image integration time to one hour period range.

While inverting these power laws could potentially provide information about the nature of the source of AR, the inversion is particularly complicated owing in large part to the dependence of frequency on the combination of the spatial scale, altitude and phase speed or evolution of the source. That particular investigation is beyond the scope of this thesis.

5.7 UNMCO Two Telescope Observations

The results from the UNMCO two telescope operations were again unexpected. Operating the two-meter separated telescopes in parallel (pointed at the same field of view) resulted in residuals that gave the impression of occasional correlations, but were generally uncorrelated. Large scale motions in the residuals (Figure 111) with timescales of about 30 minutes to over an hour gave only occasional indication of correlation Figure 112 shows an hour and 32 minute segment of residuals from a night of parallel telescope operation (large scale motions removed) with several times of apparent agreement between the two telescopes.
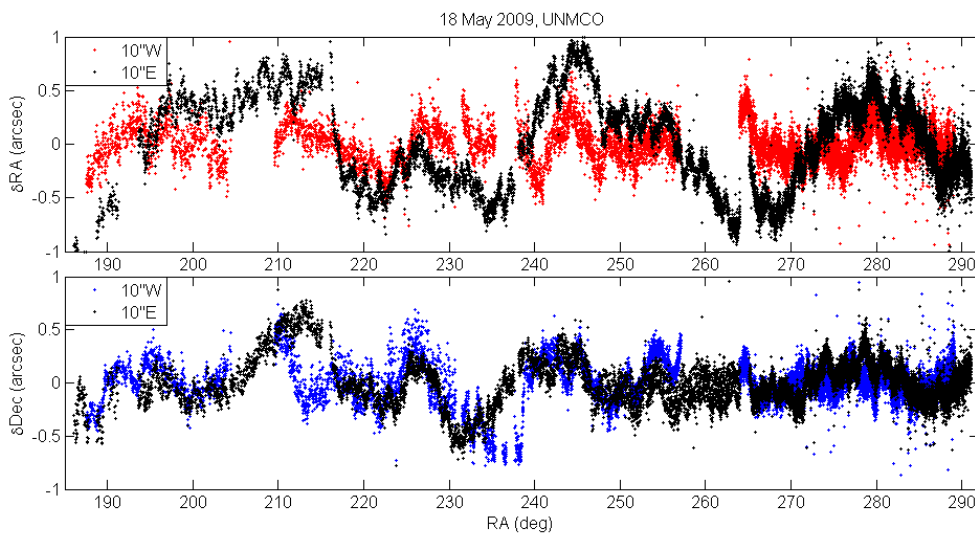
Figure 111. One night from parallel telescope operations at UNMCO.

Quantitative assessment of the level of coincidence between residuals (Figure 113) indicates no significant correlation for the night illustrated in Figure 111 and Figure 112. Additional nights of parallel telescope data confirm the lack of correlation with this telescope alignment (Figure 114 and Figure 115).
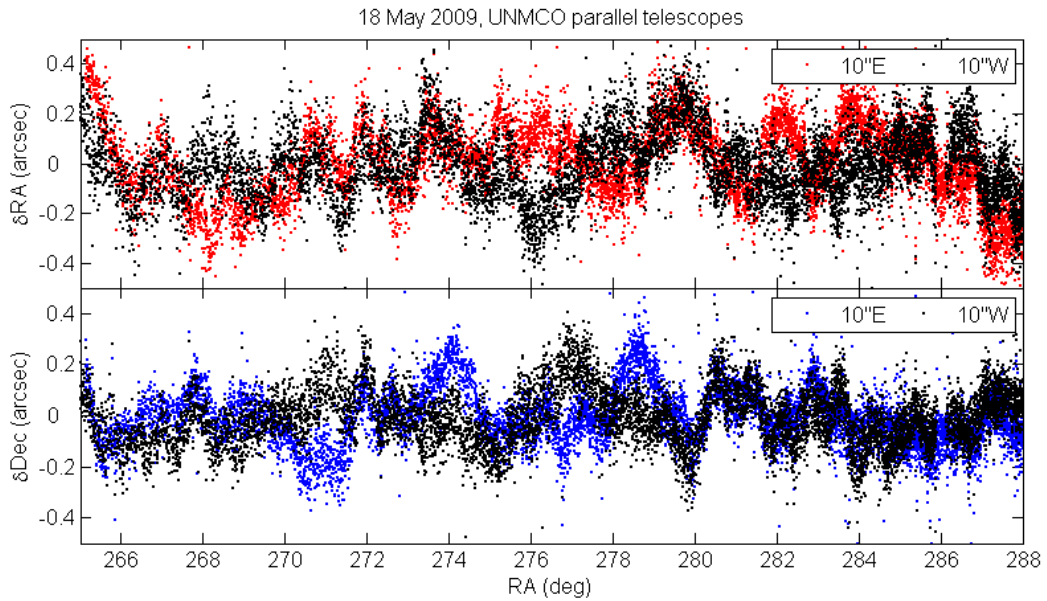
Figure 112. Section of residuals from the two UNMCO telescopes during one night of parallel operations. Motions in the residuals with timescales of order an hour or longer have been removed as in **Figure 100**.
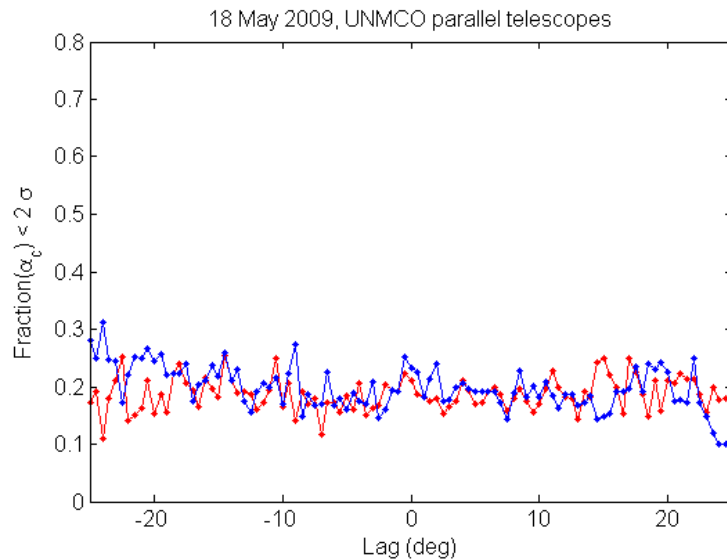
Figure 113. Fractional coincidence of residuals from the two UNMCO telescopes pointed at the same field on the sky on May 18, 2009.

202
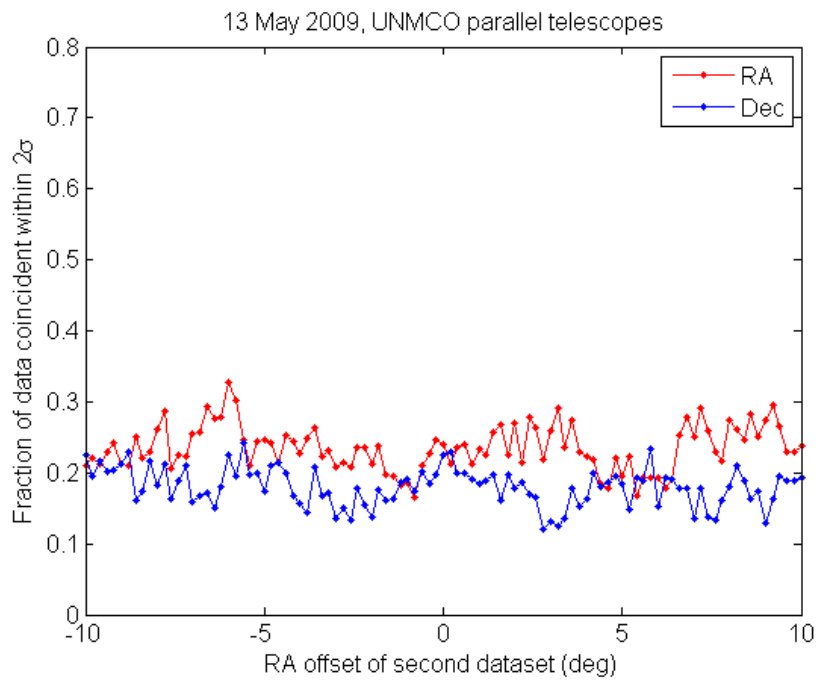
Figure 114. Fractional coincidence of residuals from the two UNMCO parallel telescopes on May 13, 2009.
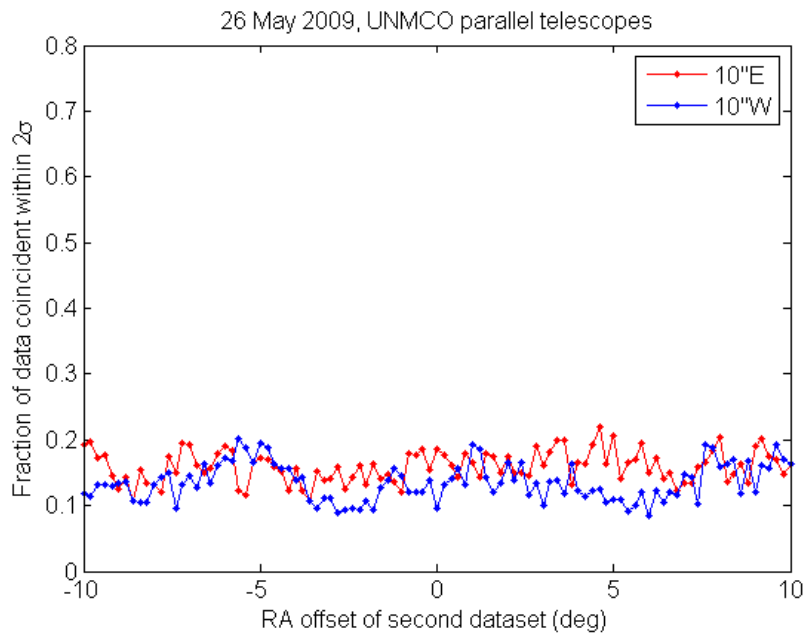
Figure 115. Same as **Figure 114** for May 26, 2009.

203

When operated in parallel, the fields-of-view of the telescopes first intersect at an altitude of about 320 meters and overlap by half of each field of view diameter at an altitude of about 640 meters. At the lower altitude each telescope is looking through an area two meters east to west by 2.8 meters north-south, such that the widest part of the atmospheric cross-section at 320 meters spans four meters. At the upper altitude, each telescope looks through an atmospheric area four meters wide for a total cross-section with a maximum horizontal span of six meters (because the fields half overlap).

Tilting the west telescope by $0°.68$ degrees towards the east lowers the initial FOV intersection to 110 meters with a maximum cross-section span of 1.4 m. The two fields-of-view fully overlap at an altitude of 170 meters and the last point of intersection occurs at 360 meters where the maximum cross-section spans 4.5 meters. One night of data obtained in this configuration indicates no change in the level of correlation of the large scale motions (Figure 116). Removing the large scale trends (Figure 117), there appears to be several periods of significant correlation between the two telescopes. This brief apparent correlation is not reflected in the overall fractional coincidence (Figure 118) indicating that it is not likely to be a significant event and that in general the residuals are uncorrelated.
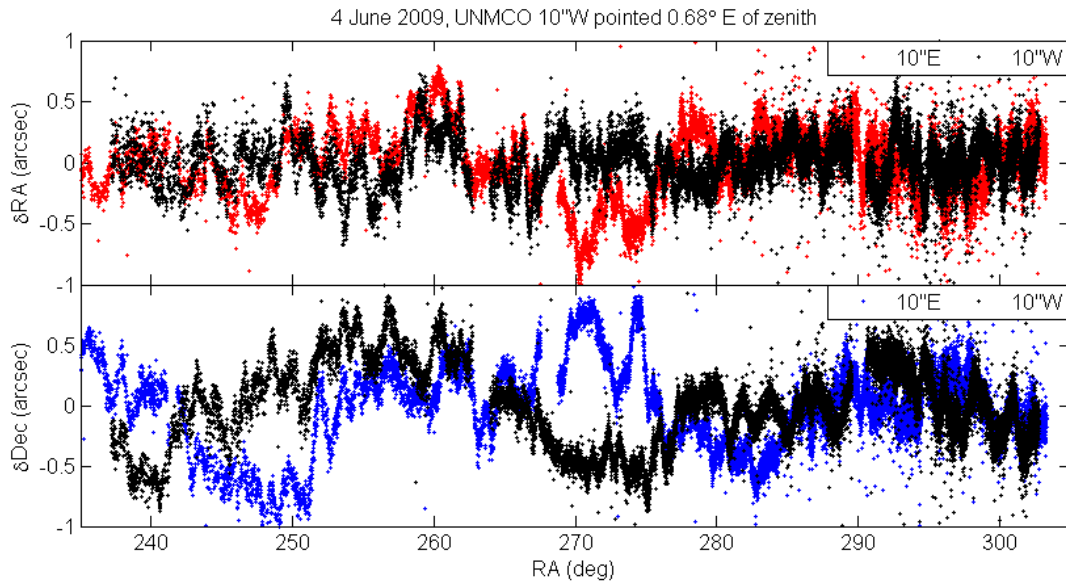
Figure 116. Comparison of residuals from June 4th 2009 at the UNMCO with the western telescope angled 0°.68 east.
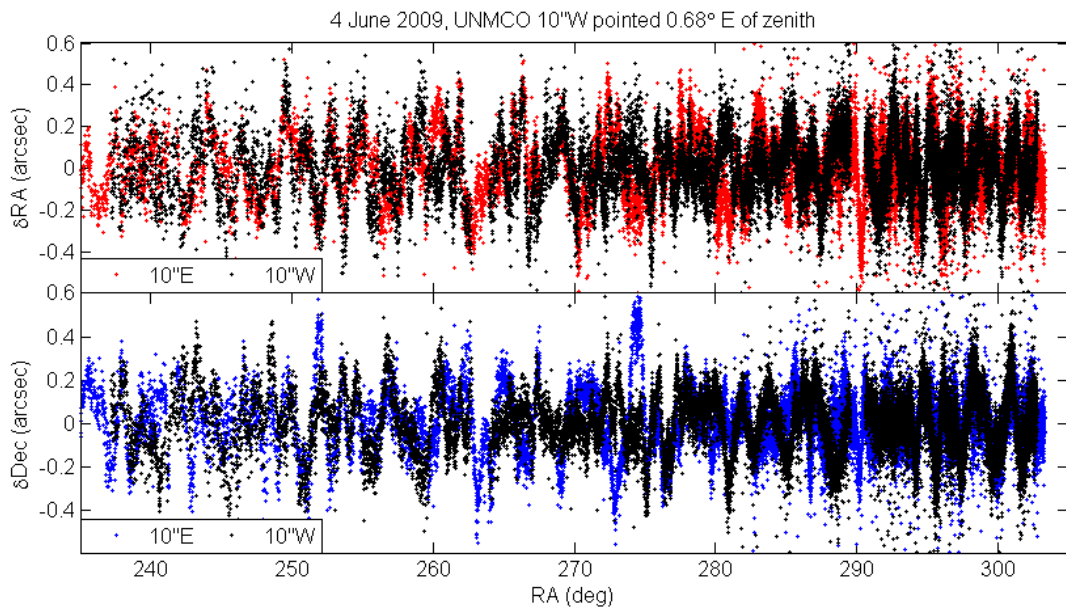
Figure 117. Same as above with large scale (30 minutes to hour plus) motions removed.

Tilting the west telescope to an angle of 1°.16 lowers the initial intersect point to just under 70 meters above the surface with a full field overlap at 100 meters and a last point of intersection at 140 meters. The large scale residuals are again unaffected by the change and remain uncorrelated. Subtracting out the large scale trends leaves residuals

that appear periodically correlated much like with the previous telescope orientations (Figure 119), although quantitative analysis again indicates no significant correlation (Figure 120).
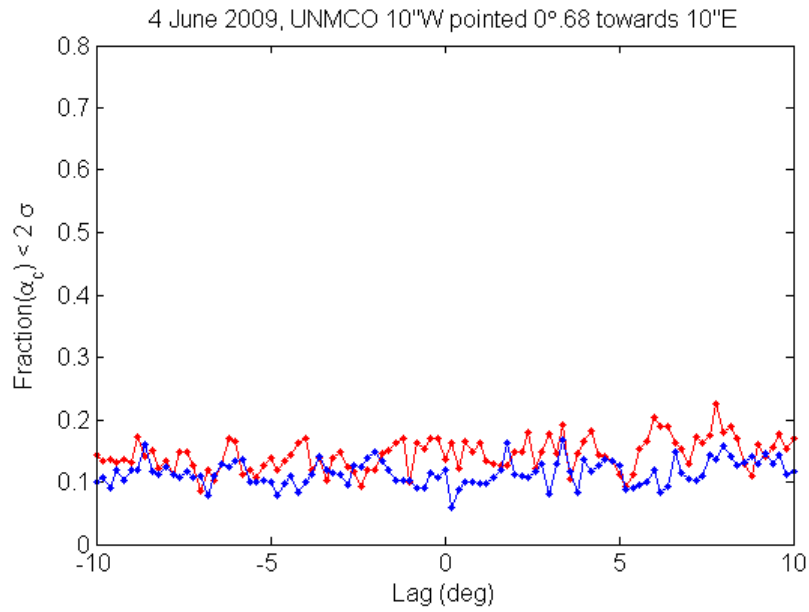
Figure 118. Fractional coincidence of residuals from UNMCO with west telescope angled towards the east telescope.
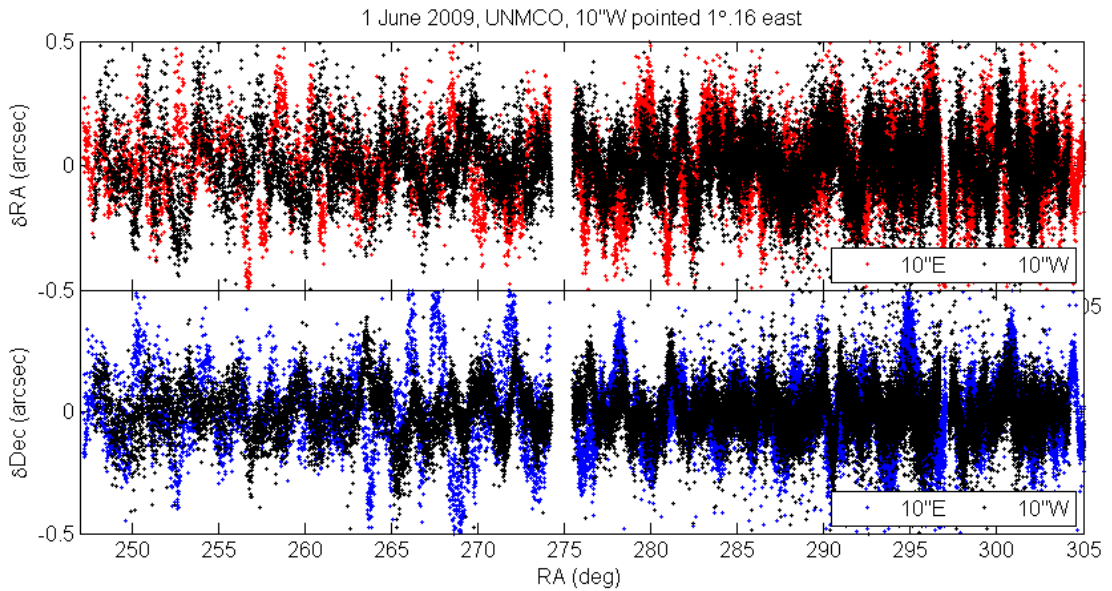
Figure 119. Comparison of residuals with 10"W pointed 1°.16 towards the east. Large scale trends have been removed.
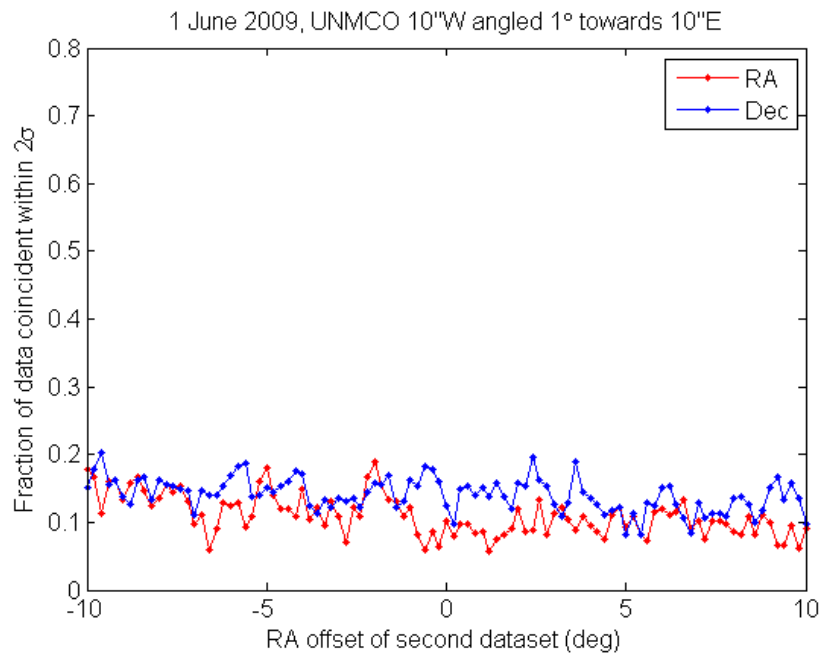
Figure 120. Same as **Figure 118** for 10"W pointed 1°.16 towards the east.

A further test involved tilting telescope 10"W away from 10"E by ten degrees. Originally, some correlation was thought to exist between parallel residuals and this experiment was conducted on the premises of determining if residuals become completely uncorrelated at large angular separations. Two nights of data were obtained with the telescopes angled apart with no clear difference between this arrangement and parallel results and no correlation (Figure 121 and Figure 122).
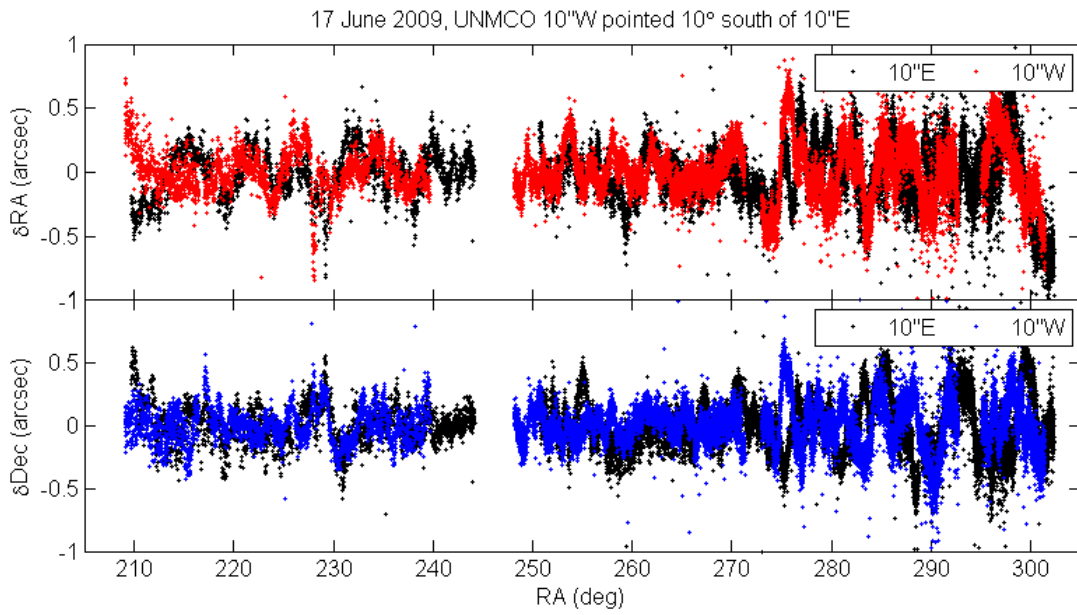
Figure 121. Residuals obtained with UNMCO 10"W pointed ten degrees south of 10"E.
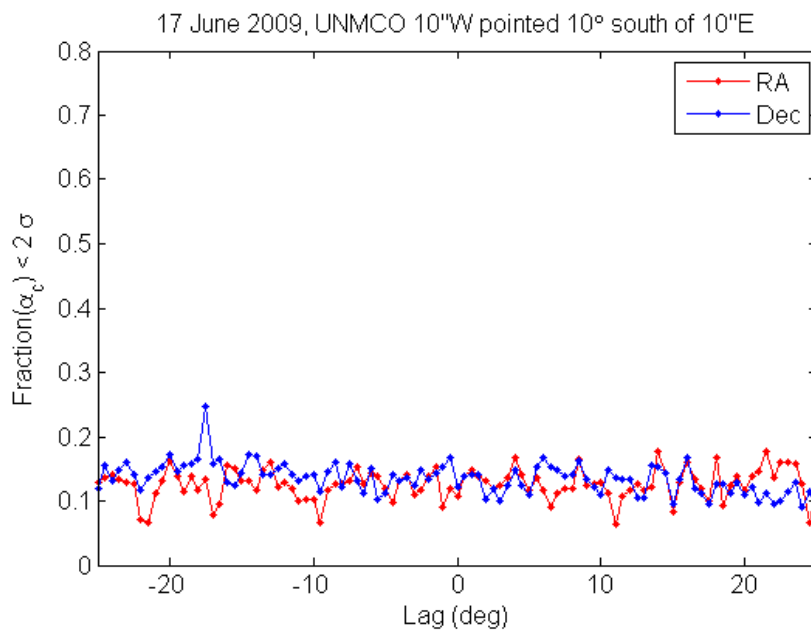
Figure 122. Fractional coincidence of residuals from widely angled telescopes.

5.8 Summary

Based on the results described above, anomalous refraction can be characterized as follows:

- A ubiquitous phenomenon that is continuously occurring at all observatory sites and for all telescopes and is characteristically consistent for all locations and instruments.

- Quasi-periodic with periods ranging from minutes to hours and amplitudes ranging from a few tenths of an arcsecond on few minute timescales to 0.5-1.5 arcseconds on tens of minutes to hour timescales.

- Power law dependences are variable from night to night, RA to Dec and telescope to telescope and range from -2/3 to -2.

- Independent of surface weather conditions, including wind speed and direction and temperature.

- Noise and residual amplitudes increased by thick (but transparent) cloud cover.

- Independent of surface pressure changes recorded by a differential microbarograph.

- Highly correlated between CCDs arrayed across a focal plane with a 2°.3 field of view, but with a consistent power-law decrease in correlation with CCD separation.

- Unrelated between telescopes separated by 50 and 300 meters observing the same field of view.

- Unrelated between telescopes separated by 2 meters, even when telescopes are angled towards or away from each other.

209

6. Conclusions and Discussion

Anomalous refraction is ubiquitous. This is a fact which bears repeating because it means that every telescope at every observatory is constantly subject to AR regardless of whether the particular observing mode used allows the AR to be seen in the data. Not only have accounts dating back to the 19th century reported observing anomalous refraction on myriad widely varied telescopes under a range of conditions, but analysis of CTI and SDSS data and observations made specifically for this research on both meter-class and 10-inch telescopes have invariably shown the continuous, characteristically consistent occurrence of AR.

This consistency between instruments and observations confirms that the effect we attribute to anomalous refraction is the same effect described in published accounts. The fact that amplitudes and periods of AR are independent of the instrument used strongly suggests that the effect is exterior to the telescope and observatory. Were AR caused by telescope structural motions, the characteristics would be expected to vary considerably between different instruments. The 2.5 meter Sloan telescope, for example, would not experience motions with the same amplitudes and frequencies as a 10-inch Meade. Further, the resonant frequencies of telescope mechanical structures are much different from those of anomalous refraction. The final and most notable point against telescope motion causing AR is the clear decrease in the level of correlation between SDSS CCDs with increasing separation between those CCDs. If AR were a mechanical effect, all CCDs in the SDSS focal plane array would move in exactly the same fashion, producing exactly the same residuals. While differences in residuals between CCDs could be

attributed to inherent differences in the CCDs (mechanical, electrical, timing, *etc.*), these differences will be independent of the CCD separation. A trend in AR correlation that specifically depends on distance between CCDs (not CCD position on the focal plane) implies an atmospheric source of limited angular extent.

Adding an additional layer to the significance of the consistency across observations, we note the broad range of enclosures housing the telescopes that have observed AR. To list just a few examples, the NOFS 1.55-m is housed in a massive dome of order ten meters in diameter, the UNMCO 10-inch telescopes are in one meter diameter domes and the Sloan 2.5-m isn't even in an enclosure at all. Once again, we can conclude that the telescope enclosure is not the likely source of anomalous refraction.

This research project was approached under the long-standing hypothesis that atmospheric gravity waves, coherent longitudinal disturbances with kilometer-scale wavelengths that propagate throughout the atmosphere, were the source of anomalous refraction. Numerical simulations of the refraction effects caused by AGWs indicated that the temperature gradients associated with waves within a few kilometers above the surface are capable of causing refractions with amplitudes comparable to those of AR. The few minutes to hour timescales associated with propagating AGWs were also well matched to the observed periods of AR.

Based on this hypothesis, we devised an experiment wherein parallel pointing telescopes with sub-kilometer spacing would simultaneously make astrometric observations through

the same wave structures. Because AGWs are coherent, non-stochastic phenomena, any particular wave structure is assumed to maintain a quasi-uniform period and amplitude for at least a full wavelength. Therefore, the parallel telescopes with sub-wavelength separation should experience anomalous refraction that is highly correlated the majority of the time, assuming the AR is caused by kilometer-scale waves.

The results of performing this experiment at NOFS showed that parallel pointed telescopes separated by 50 to 300 meters observed anomalous refraction that showed no correlation. This result was consistent over four nights and three telescopes, during which time anomalous refraction was observed to occur continuously. Anomalous refraction is therefore conclusively shown to not be caused by kilometer-scale atmospheric gravity waves at any altitude. The source of anomalous refraction is additionally constrained by these results to be either non-propagating or rapidly changing such that period and amplitude characteristics are considerably modified within the space of 50 meters (the closest spacing of the NOFS telescopes). In either case the scale over which the source of anomalous refraction is coherent must be smaller than 50 meters.

The AR observed by the SDSS was highly coherent across all CCDs in the 2°.3 focal plane, showed no phase lag that would indicate a travelling disturbance, but did show slight modulation in the AR between widely separated CCDs. To understand this combination of features, the geometry of the fields viewed by the individual CCDs and how they intersect as a function of altitude must be considered (Figure 123). Immediately above the telescope, all of the CCDs are looking through the same volume

of air. By an altitude of approximately 75 meters the outermost r CCDs are no longer looking through the same column of air, while the fields of view of r_1 and r_5 diverge at an altitude of about 100 meters. The fields of all of the CCDs are separate above 800 meters. One explanation for the high degree of correlation across the field is a non-propagating disturbance above the divergence altitude of most of the CCDs, with a coherence scale larger than the SDSS FOV, but small enough that the phase change across the FOV results in an observable change in refraction. Another possibility is a propagating source at very low altitude (*i.e.* below about 50 m), which would be seen simultaneously by all 30 CCDs and thus should produce no phase lag. If the disturbance had a coherence scale slightly smaller than the combined FOV, the CCDs would still be highly correlated, but the level of correlation would decrease with increased angle between the individual fields as is seen in the results. The wavelike temporal structure of the anomalous refraction would then be explained either by modulation of the disturbance with time or a wave-like structure drifting over the telescope.

It should be noted in discussing SDSS phase lags that all SDSS residuals are frame averages, not star-by-star offsets. In the NOFS and UNMCO data it is possible to see sub-integration time residual motions, but the frame-averaging of SDSS removes any of these signatures. Because of this, the resolution to which we can observe phase-lags is limited to a frame-length and any wave phase that crosses the focal plane within that time (36 seconds) will effectively appear to encounter all CCDs simultaneously. Adding to the possible sources of the observed SDSS residuals is a large-scale high altitude

disturbance that propagates quickly enough for a single phase to cross the entire FOV in less than 35 seconds.
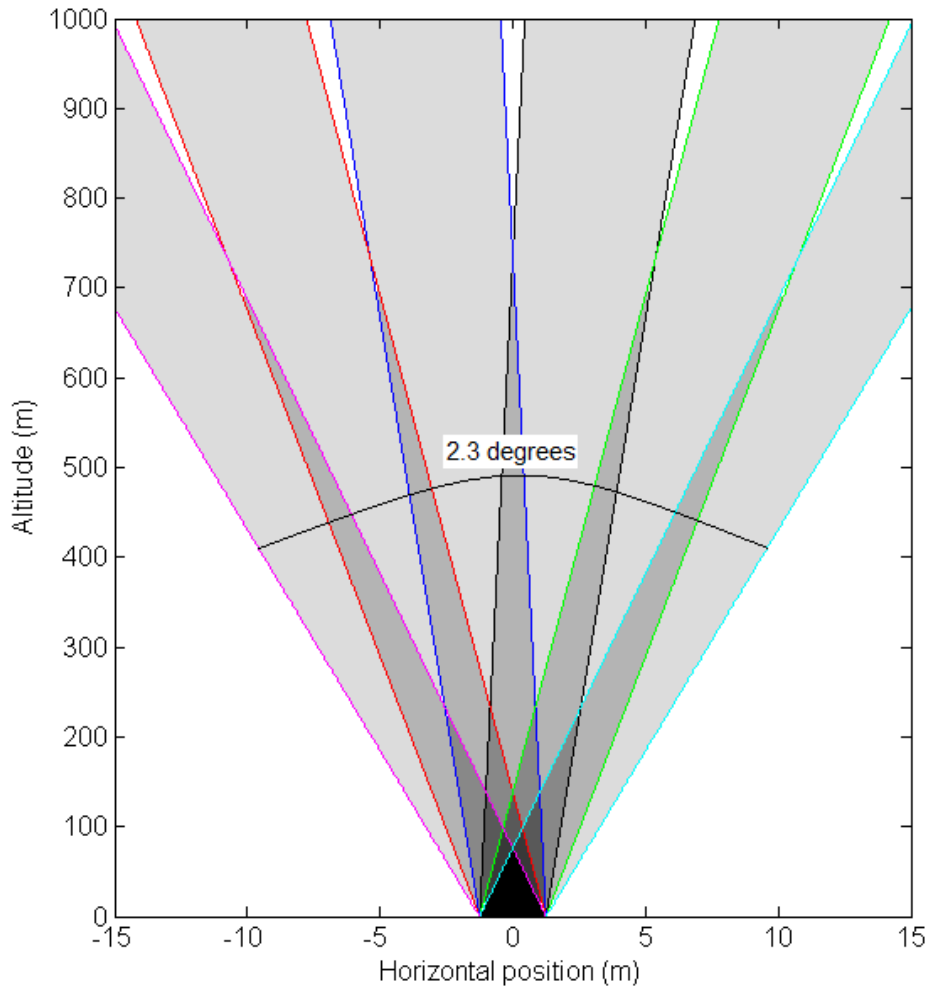
Figure 123. Fields viewed by each of the six SDSS *r'* CCDs (arrayed in a row perpendicular to the scan direction. Shading indicates the number of *r'* CCDs looking through a particular part of the atmosphere (from 0 – white, to all 6 – black). Each CCD has a field of view of 13.65 arcminutes, with a center-to-center CCD separation of 25.2 arcminutes across columns and 18 arcminutes down columns (see **Figure 30**)

Another explanation for the SDSS behavior is very short wavelength waves propagating fast enough that multiple wavelengths cross all fields of view in the timespan of a single exposure. The total amplitude of the AR due to this source would be the average tilt over the field of view, while the variation across the FOV would be explained by different

CCDs seeing a slightly different overall phase. There would be no phase lag as the field-crossing time for a single wave would be shorter than the exposure time.

The problem with this latter hypothesis is that the amplitude of the observed refraction would be inversely proportional to the ratio of the aliased wave frequency to the exposure time, and as a rule much smaller than the refraction amplitude of the unaliased wave. Shorter exposure times would result in larger refractions and astronomers using exposure times much shorter than the wave period would see very large amplitude refractions. Beyond wave periods a few times smaller than the exposure length, the aliased refraction would become unobservable. For the SDSS exposure time of 54 seconds, an observed (aliased) AR of a few tenths of an arcsecond could be caused by a 20 second period wave with an amplitude of ten arcseconds. A three second period, ten arcsecond wave would produce an aliased refraction of less than 0.01 arcseconds. A 20 second wave with a one arcsecond amplitude would at most produce a refraction of 0.03 arcseconds. Because astronomers using very short exposure methods do not see refractions of tens of arcseconds on ten second timescales (nor did astrometrists making "eyeballed" meridian circle measurements 100 years ago), we rule out high frequency aliased waves as a source of anomalous refraction.

Combining the results of the SDSS data with the results from NOFS provides a further constraint on the source of anomalous refraction. The highly correlated nature of the SDSS residuals gives us a lower limit on the angular scales of AR, while the generally uncorrelated nature of the NOFS residuals provides a maximum spatial scale. By

combining these two limits, an approximate upper altitude boundary for the source of AR can be derived. Placing an imaginary Sloan telescope between the NOFS 1.3-m and 1.55-m telescopes, it is apparent that the source of AR must occur below the altitude where the Sloan FOV overlaps with the fields-of-view of the two NOFS telescopes (Figure 124). Above this altitude, a disturbance large enough to be coherent across the SDSS FOV would also be coherent between the two closest NOFS telescopes. The three fields of view will initially overlap slightly above 1000 meters and the outer edges of the Sloan FOV will cross the centers of the two NOFS telescopes' fields of view at approximately 1250 meters. The source of anomalous refraction must therefore occur below approximately 1000 meters.
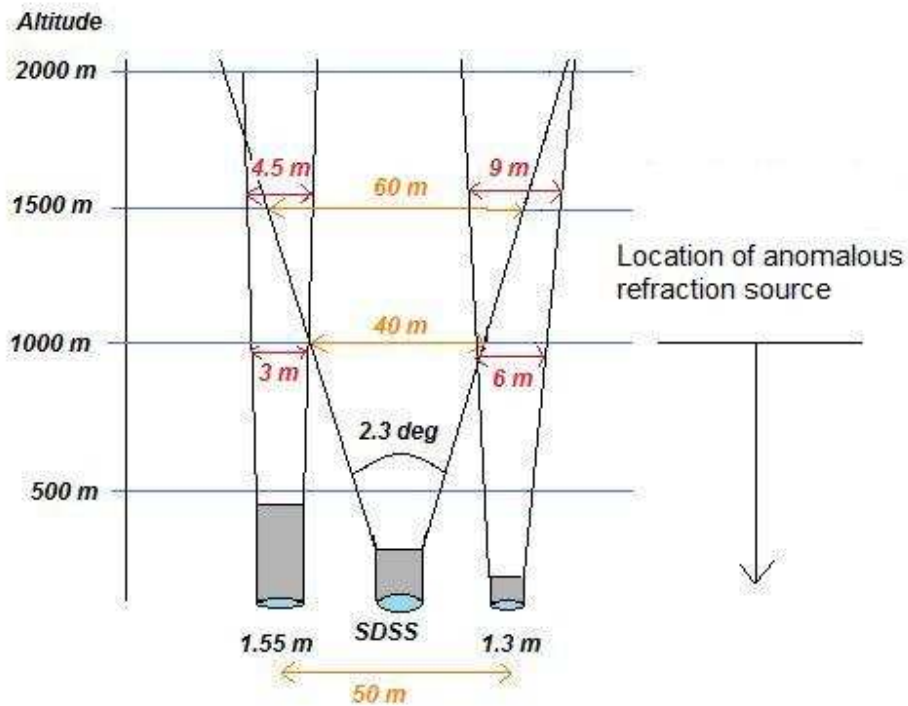
Figure 124. Constraint on the maximum altitude of the source of AR. Placing an imaginary Sloan telescope between the two closest NOFS telescopes and calculating the altitude at which the fields of view overlap places an upper limit on the altitude of the AR source.

216

The main caveat to this argument is we're assuming that the source of anomalous refraction is the same (characteristically) at all observatories. This is not an unrealistic assumption considering that the characteristics of AR are consistent across observatories. However, there may be minor differences in spatial scale or altitude, for example, between the observatories that are not immediately clear in the data.

The observations made at the UNM Campus Observatory were designed to further constrain the scales (both spatial and angular) and altitudes of the source of AR. The two telescopes are arrayed roughly east-west with a separation of about two meters. Both have fields of view in the east-west direction of 0°.36, which places the initial FOV overlap at an altitude of about 320 meters. In order for AR to be correlated between these two telescopes, the source must either occur above this altitude or below this altitude with a spatial coherence scale that is larger than two meters. In order for the AR to be uncorrelated between the telescopes, the source must exist below the overlap altitude and must have a coherence scale of less than two meters.

An AR source existing above the overlap altitude could potentially produce uncorrelated AR if the source was sufficiently incoherent on scales smaller than the combined FOV, such that the lack of correlation between the non-overlapping wings of the fields of view dominated the overall AR. This becomes increasingly unlikely with increased fractional field overlap, *i.e.* with increasing altitude (Figure 125). In the range of altitudes where this might be considered a possibility, *i.e.* from about 320 m to about 800 m, only the immediately adjacent CCDs on the SDSS overlap and the combined FOV of the Sloan

telescope has a cross-sectional area many times larger than that of the two UNMCO telescopes. An AR source with coherence scales smaller than the UNMCO combined FOV at these altitudes would be distinctly incoherent across the CCDs of the Sloan telescope, indicating that the uncorrelated residuals of the UNMCO telescopes must be due to a source of AR that is located below the initial overlap of the two telescopes.
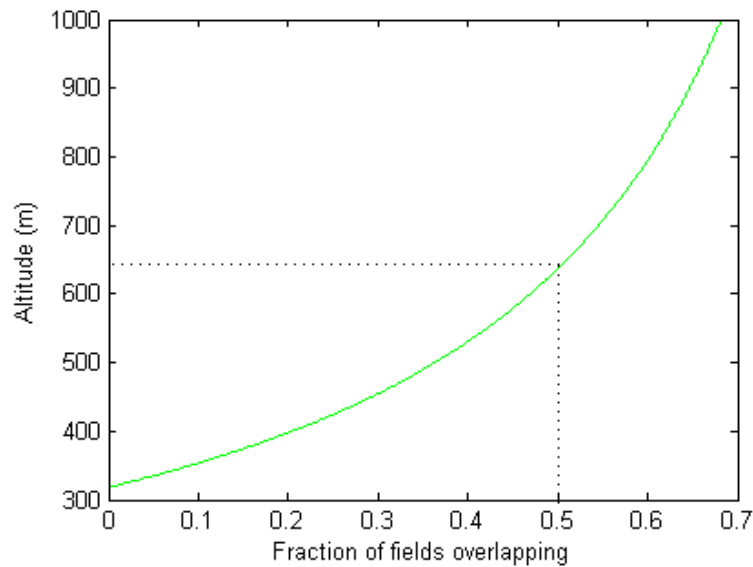
Figure 125. Fractional overlap of UNMCO 10" telescope fields of view as a function of altitude. The fields overlap by half at an altitude of about 650 meters.

We have observed that in the parallel configuration the AR seen by the two UNMCO telescopes is completely uncorrelated. Based on this observation and the above arguments, we can immediately confine the primary source of AR to an altitude below 320 meters and spatial coherence scales of less than two meters. Additionally, the refractive structure of any source that is propagating must completely change in the space of two meters, except under the very unlikely condition that the AR source never propagates in the east-west direction (*i.e.* never passes over first one telescope then the other).

218

A possible atmospheric explanation is a microscale disturbance such as a train of meter-sized Kelvin-Helmholtz billows as described by Chimonas (1999, see Figure 20). The weakness of the microscale K-H billow model is the very small size requires these disturbances to be either stationary or very slowly propagating, otherwise the periods become much shorter than the integration time and we return to the aliasing argument. Wavelike variations in refraction would result from the changing structure of the billows and growth of instability with time.

A numerical analysis of the refraction due to a non-propagating microscale K-H billow (Figure 126) lends credence to this possibility despite the drawbacks. The wave is simulated in the manner described in Section 2.3.3 and modeled with a 3-m billow size and 0.1 Kelvin temperature change across the disturbance. Over the course of 30 minutes initial waves grow, crest, curl-up and then dissolve into turbulence. The resulting refraction (calculated using Equations (19) through (22)) varies considerably in both time and space with a telescope located two meters down the wave line seeing dramatically different refractive structures. The refraction traces are convolved with a 100 second boxcar filter to simulate the filtering effects of UNMCO exposure-time integration of positional offsets.
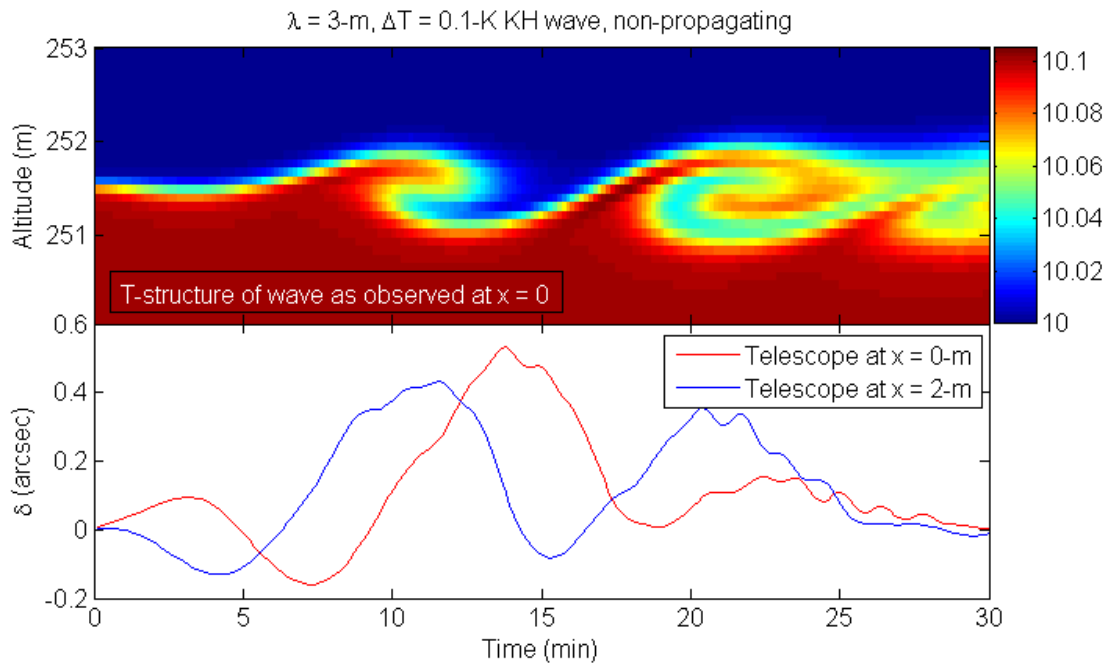
Figure 126. Simulated stationary microscale Kelvin-Helmholtz instability and resulting refraction seen by two telescopes separated by two meters parallel to the wave train. The wave structure in the top plot is the one-dimensional temperature field as seen by an observer at located at x = 0 as a function of time.

Because K-H waves are primarily caused by wind shear, they can be expected to propagate with the mean wind at the level where they are created. Typical nocturnal winds increase from often calm conditions just above the surface to a maximum of anywhere from 10 m/s to 30 m/s near the top of the boundary-layer (which ranges from 100 to 500 meters above the surface, with typical values of 200 to 300 meters: Figure 127, Stull 1988). A two-meter wavelength K-H wave at an altitude of 300 meters propagating at 15 m/s would have a frequency of 7.5 Hz and cross the fields of both telescopes in less than half a second. If K-H waves at the surface are the source, we would expect a direct relationship between wave periods and surface wind speeds, which is distinctly not the case.
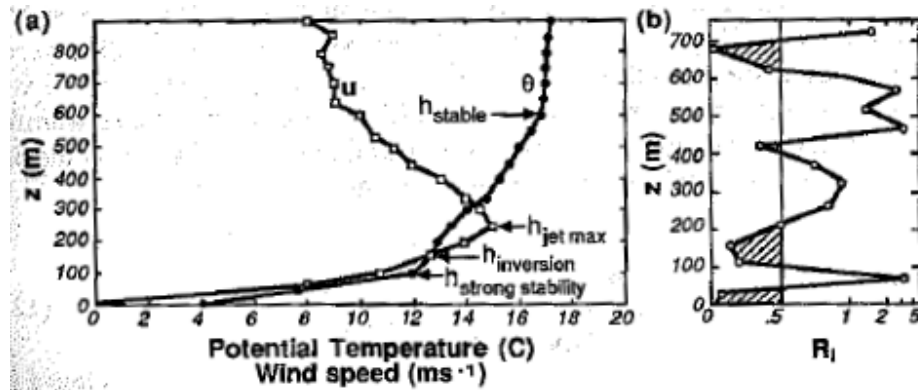
Figure 127. Example of nocturnal boundary layer wind and potential temperature profiles (a), and the corresponding Richardson number profile (b). (From Stull 1988)

The complex layered nature of the atmospheric boundary-layer, with various waves, turbulent patches and other disturbances coexisting and feeding off of each other at all times, lends credence to the idea that anomalous refraction could be caused by a superposition of multiple sources at multiple altitudes. Figure 126 indicates that very small-scale disturbances with associated temperature fluctuations of tenths of Kelvin can create refractions as large as several tenths of an arcsecond.

Additional experiments at UNMCO involved tilting the telescopes towards or away from each other. Tilting the telescopes away from each other was intended to pin down the outer angular scales of AR, *i.e.* to find the angle at which the AR seen by the two telescopes no longer showed any correlation (early analysis of parallel telescope data suggested occasional correlation between the residuals, which has since been disproved). Angling the western telescope towards the east lowered the altitude at which the fields of view overlapped, but also decreased the total volume of atmosphere containing the intersection of the two fields. The goal of this latter experiment was to determine if

lowering the initial overlap altitude below the primary altitude of the source of AR would result in the telescopes seeing primarily correlated residuals.

Tilting the telescopes away from each other by up to $1°.5$ increased the angular size of the area viewed to as large as $1°.86$, but did not change the correlation characteristics of the residuals. Tilting the telescopes apart by $10°$ again had no affect on the relationship between the residuals. After determining that the residuals from the parallel observations were uncorrelated, these tests became unnecessary because it was never hypothesized that residuals would become more correlated with increasing angular separations and with no correlation to begin with, they couldn't become less correlated. Tilting the telescopes towards each other lowered the initial overlap altitude to first 170 meters ($0°.68$ tilt, Figure 116) and then 70 meters ($1°.16$ tilt, Figure 119), but did not produce any correlation between the two telescopes.

The primary complication with the "crossed-beams" experiment is that although the region of the atmosphere where the fields cross can be made very low, much larger fractions of the air columns through which the telescopes are looking are completely unrelated. This is a similar situation to the CCDs in the Sloan focal plane, with the large exception that the Sloan CCDs look through the same air volume from the surface up to an altitude of 75 meters.

The lack of correlation between the UNMCO telescopes places an outer limit on the scales over which the source of anomalous refraction are correlated at two meters. This

poses the problem when considering the SDSS results that the SDSS aperture is 2.5 meters in diameter and all SDSS CCDs see approximately the same refraction. In order for both SDSS and UNMCO to have their observed correlated and non-correlated refractions, respectively, the source of anomalous refraction must be located in the lowest few-hundred meters of the atmosphere where the fields viewed by all SDSS CCDs overlap (see Figure 128).
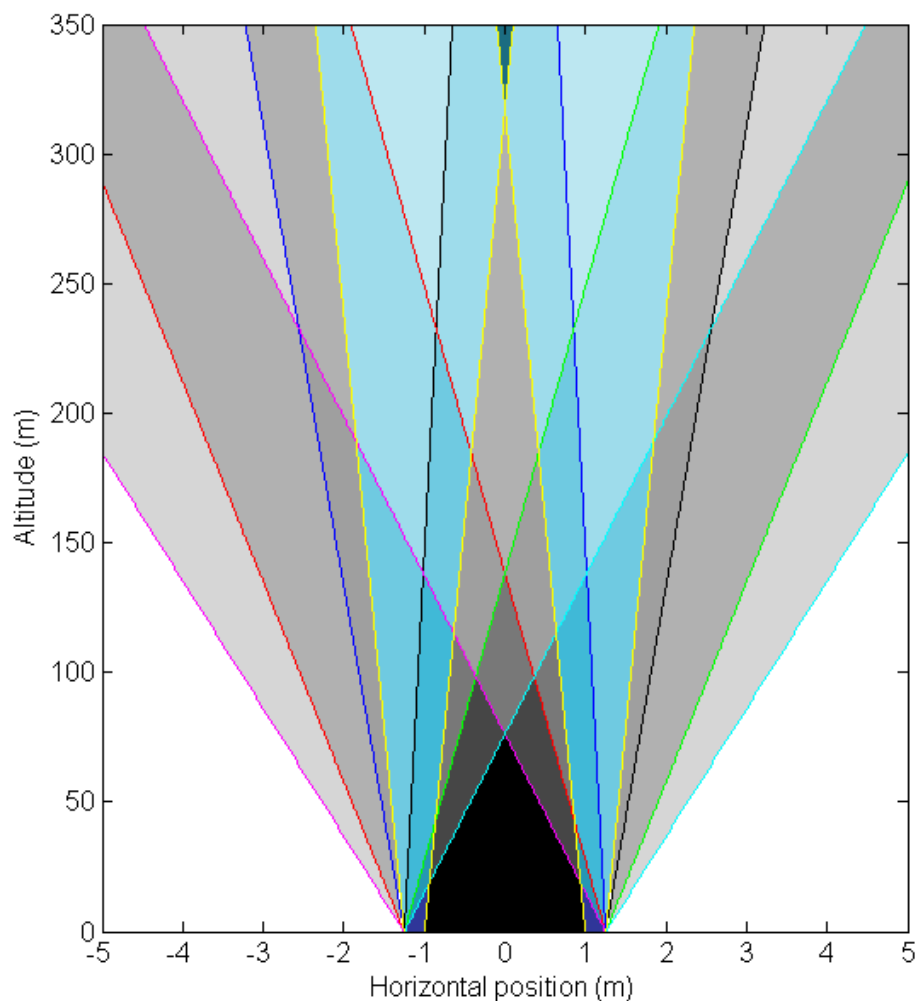
Figure 128. Fields viewed by the two UNMCO telescopes (blue) overlaid on the fields viewed by the first row of SDSS CCDs with number of overlapping CCDs indicated by shading. All CCDs see the same volume of air in the black triangle at the base.

The primary constraints we have placed on the source of anomalous refraction based on astrometric observations are that it is very low altitude (up to a few tens of meters above the surface) constantly occurring, and coherent across scales of order two meters or less. The source is very slow moving, so as to produce periods of minutes to tens of minutes and no observed phase lag between CCDs and telescopes.

The microbarograph data showed no relationship with the astrometric data. Because the differential microbarographs were designed to detect boundary-layer waves hundreds of meters to kilometers in wavelength, they are incapable of detecting meter scale phenomena. The approximately ten meter baselines of μBAR would only see aliased differences in atmospheric pressure. Typical surface pressure fluctuations due to kilometer scale waves with amplitudes of tens to hundreds of meters are in range of microbars to hundreds of microbars, with the pressure differences across ten meters being considerably smaller. If the source of AR were AGWs of a few meter wavelengths (were such a phenomenon to exist, a fact not supported by the literature), assuming the same amplitude to wavelength ratios applied to a microscale AGWs, then from Equation (35) the surface pressure perturbations would be three orders of magnitude smaller than for the kilometer scale waves. Even a point source microbarograph (with a resolution of order a microbar) would be incapable of seeing a surface pressure change of 0.001 to 0.1 microbars. So while the microbarographs add additional proof that AR is not caused by AGWs, they cease to be useful in the detection of meter-sized atmospheric phenomena.

The weather station data indicated that not only does AR occur regardless of the weather conditions, but the characteristics of AR are independent of the weather (and particularly wind) conditions. Because the observed frequencies of AR do not change with wind speed, the source of AR must not be moving with the mean wind. Shear initiated phenomena such as K-H waves tend to travel with the mean wind at the level where they are created. An exception might be if the shear disturbance were caused by the motion of air over a fixed object (*e.g.* buildings, trees, *etc.*), but even then, the source of shear would disappear under calm conditions and AR does not do likewise.

A wide array of additional experiments exist that will further constrain the source of anomalous refraction and potentially pin down what that specific source is (or sources are). Further studies of the SDSS dataset will involve examining the cross-focal plane correlation of runs as a function of zenith angle. Additional campus observatory operation could including comparing residuals obtained with the telescope dome removed, verses those with a telescope dome, potentially indicating if the source of AR is directly influenced by surface conditions. Physically moving one of the UNMCO telescopes as close to the other telescope as possible would allow maximum overlap of the two fields of view. If the two telescopes were separated by half a meter and arrayed north-south, the fields would initially overlap at an altitude of 57 meters. This arrangement has the advantage over the crossed-beams that the overlap increases with altitude. If the source of AR occurs above 60 meters it would appear primarily correlated in this configuration. A very small tilt of one or both telescopes would lower this overlap altitude while maintaining the overall overlap at higher altitudes.

To pin down the specific atmospheric source of AR, we require additional atmospheric instrumentation. The best means of studying boundary-layer phenomena in the lowest hundred meters of the atmosphere is with instrumented meteorological tower (Figure 129). A tower containing numerous levels of thermistors (for temperature measurements) and anemometers (measuring horizontal wind speed and direction and vertical wind speed) would allow us to pin down the specific altitude (or altitudes) below 100 meters at which the source of AR was occurring. We would also be able to determine whether the source was a wave structure or more turbulent in nature and the stability of the atmosphere, which would indicate whether buoyancy, convective or shear phenomena could exist. The cost and logistics of such an installation are prohibitive for the UNMCO; however, future research at an established observatory, or even the creation of a specific site for the detailed study of the atmosphere and astronomy would prove immensely valuable for both fields of study.

The possibility remains wide open that the source of anomalous refraction is a phenomenon as yet unfamiliar to atmospheric physicists. The unconventional method of observing the nocturnal boundary-layer using astrometry may be sensitive to dynamical phenomena not seen with more traditional atmospheric instruments. Whether or not this is the case, by pinpointing this mysterious phenomenon, we not only make a vital contribution to ground-based astrometric observing, but also to the study of the nocturnal boundary-layer.
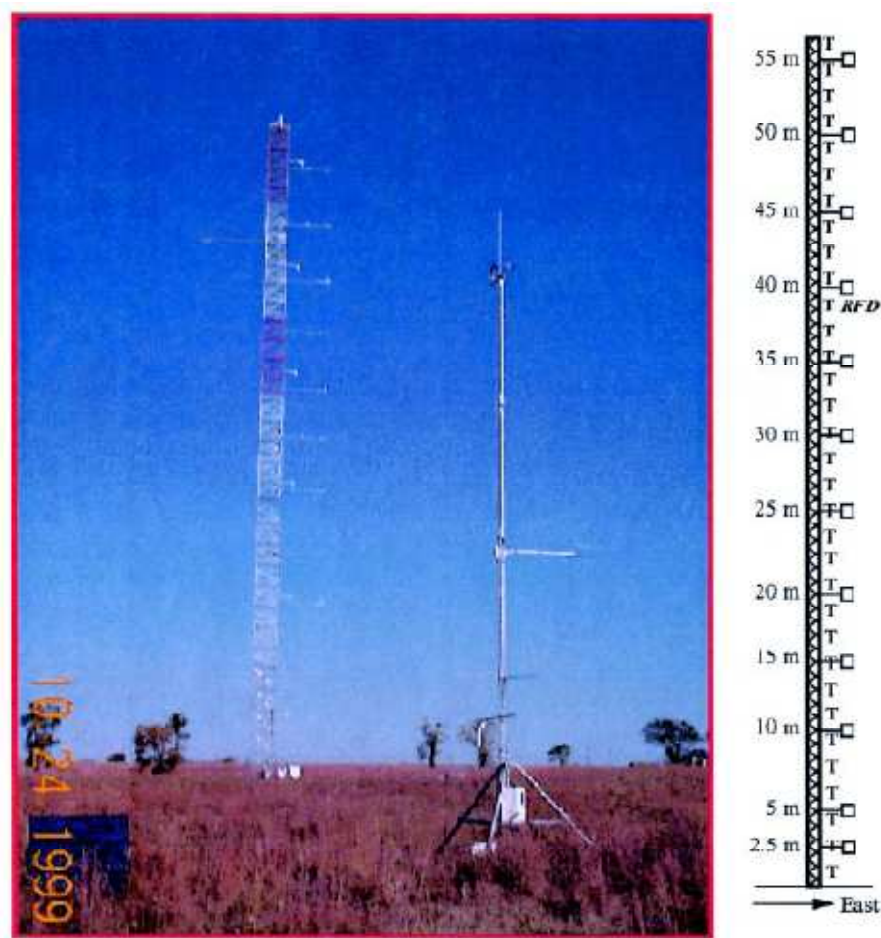
Figure 129. 60 meter meteorological tower (left of center in image) used by the CASES-99 atmospheric research campaign. The tower was heavily instrumented with weather stations, anemometers and temperature probes as indicated in the schematic at right (from Poulos *et al.* 2002)

The scales of AR are clearly small enough that future very large-aperture instruments such as LSST will suffer significant time-variable astrometric distortion within the FOV and the ubiquitous occurrence of AR will make this an issue of constant concern. Considering the decrease in correlation with large CCD separations in the 2°.3 Sloan FOV, far less correlation can be expected across the 3°.5 diameter, 189 CCD LSST field. LSST will image the sky taking 15 second snapshots of each field and stitching them together to create an all-sky mosaic (Ivezic *et al.* 2008). Each 15 second snapshot will

227

contain AR distortion and the AR will change significantly over the course of several exposures causing distortion between each frame of the mosaic. Through understanding the source of AR, we may find methods akin to adaptive optics for mitigating the effect, or an intelligent solution for removing the positional errors in processing.

The source of AR is potentially a major player in the energy budget and dynamics of the nocturnal boundary-layer and ground-based astrometric data may prove an invaluable new source of data on this phenomenon. Understanding the atmosphere is vital to maintaining safety and quality of life on the local level (through weather forecasting, storm monitoring, *etc.*) and for the health of the planet on the global level (through studies of climate change). Any new source of information on the nature of the atmosphere should not be taken lightly.

List of Appendices

229

Appendix A *refractivity.m*

%refractivity.m - calculate refractivity of moist air given temperature, pressure,
% humidity, etc.

function Mu = refractivity(p, T, wv, lambda)

%inputs: pressure (Pa)
%   Temperature (K)
%   mixing ratio of water vapor (g/kg)
%   wavelength of light (microns)
%output: Refractivity of moist air

%Standard atmosphere
Ts = 288.15; %standard temp (15C)
Ps = 101325; %standard press (Pa)
xc = 450; %ppm C02, standard
Mw = 0.018015; %kg/mol, molar mass of water vapor
R = 8.314510; %J/mol*K, Gas constant

%Constants for standard refractivities of dry air
k0 = 238.0185; %1/micron^2
k1 = 5792105; %1/micron^2
k2 = 57.362; %1/micron^2
k3 = 167917; %1/micron^2

%Constants for standard refractivities of water vapor
w0 = 295.235; %1/micron^2
w1 = 2.6422; %1/micron^2
w2 = -0.032380; %1/micron^4
w3 = 0.004028; %1/micron^6

%Constants for density equation

%For saturation vapor pressure of water
A = 1.2378847e-5; %1/K^2
B = -1.9121316e-2; %1/K
C = 33.93711047;
D = -6.3431645e3; %K

%For enhancement factor of water vapor
al = 1.00062;
be = 3.14e-8; %1/Pa
ga = 5.6e-7; %1/degC^2

%For compressibility
a0 = 1.58123e-6; %K/Pa
a1 = -2.9331e-8; %1/Pa
a2 = 1.1043e-10; %1/K*Pa
b0 = 5.707e-6; %K/Pa
b1 = -2.051e-8; %1/Pa
c0 = 1.9898e-4; %K/Pa
c1 = -2.376e-6; %1/Pa
d = 1.83e-11; %K^2/Pa^2
e = -0.765e-8; %K^2/Pa^2

%Convert temperature to Celsius
t = T - 273.15;

%Use variable precision accuracy: vpa (32 digits)

%Enhancement factor of water vapor in air
f = (al + be*p + ga*t^2);

%Partial pressure of water vapor
pw = (p/(0.62197*wv/1000+1)); %Pa

%Molar fraction of water vapor in moist air
xw = vpa(f*pw/p);

%Refractivity of standard air, no humiditiy
n_as = vpa((k1/(k0-(1/lambda)^2) + k3/(k2-(1/lambda)^2))*10^-8 + 1);

%Effect of CO2 concentration
n_axs = vpa((n_as-1)*(1 + 0.534e-6*(xc-450))+1);

%Refraction of standard water vapor
n_ws = vpa((1.022*(w0 + w1*(1/lambda)^2 + w2*(1/lambda)^4 +
w3*(1/lambda)^6))*10^-8 + 1);

%Molar mass of dry air
Ma = (28.9635 + 12.011e-6*(xc-400))*10^-3; %kg/mol

%Convert standard temperature to Celsius
ts = Ts - 273.15;

%Compressibility of dry air
Za = 1 - (Ps/Ts)*(a0 + a1*ts + a2*ts^2) + (Ps/Ts)^2*(d);

%Constants for calculation of water vaper compressibility
Pw = 1333; %Pa water vapor pressure

Tw = 293.15; %K water vapor temperature
tw = Tw - 273.15; %Convert T to Celsius

%Compressibility of water vapor
Zw = 1 - (Pw/Tw)*(a0 + a1*tw + a2*tw^2 + (b0 + b1*tw) + (c0 + c1*tw)) +
(Pw/Tw)^2*(d + e);

%Density of standard air
Rhoaxs = (Ps*Ma/(Za*R*Ts));

%Density of water vapor
Rhows = (Pw*Mw/(Zw*R*Tw));

%Compressibility of moist air under experimental conditions
Z = 1 - (p/T)*(a0 + a1*t + a2*t^2 + (b0 + b1*t)*xw + (c0 + c1*t)*xw^2) + (p/T)^2*(d +
e*xw^2);

%Density of dry component of moist air
Rhoa = p*Ma*(1-xw)/(Z*R*T);

%Density of water vapor component of moist air
Rhow = p*Mw*xw/(Z*R*T);

%Index of refraction of moist air
n = vpa(((Rhoa/Rhoaxs)*(n_axs - 1) + (Rhow/Rhows)*(n_ws - 1)) + 1);

%Refractivity of moist air
Mu = double(n) - 1;

Appendix B *getrowtime.m*

%getrowtime - retrieve timestamp of a row and calculate time of row dump.

function [time] = getrowtime(lsb, msb, row)

%input: vectors containing least significant time word (1.0m col 3),
%   most significant time word (column 4) and row number
%output: Row time in fractional UT hours

%Timing data are contained (along with other telescope data) in
%the non-science columns preceeding the image columns (underscan).
%For the 1.0 meter telescope, columns 3 and 4 contain the time stamps.
%Time is recorded in units of integer ticks with the precision of the
%time record determined by the number of ticks per second, which is
%dependent on the computer operating system used. For the 1.0 meter
%there are 10000 ticks per second - providing 0.1 ms precision
%
%Time is recorded as two parts of a 16 bit number
%The most significant word is recorded as multiples of 2^16
%The least significant word ranges from 1 to 2^16
%The time (sec) is calculated by adding the most sig. word X 2^16
%to the least sig. word and multiplying the result by 0.0001
%--%

%least significant word is a 16 bit number recorded in 15 bits -
%number cycles from 1-2^15 then jumps to -2^15 and cycles back
%to zero.

%If number is < 0, add 2^16 to get actual 16 bit number.
if lsb(row) < 0
    lsbc = lsb(row)+2^16;
else
    lsbc = lsb(row);
end

ticks = msb(row)*(2^16)+lsbc; %Add two words to get number of ticks

time = (ticks*1e-4)/3600; %Convert ticks to hours

Appendix C *prep13_1.m*

```
%prep13_1 - copy 1.3m frames to new files and change header

%Program creates a list of all raw frames, calculates epoch, central row
%meridian crossing time, and LST and inputs them in headers of fits files
%copied to proc directory.
%Calls matlab programs getrowtime.m, julday1.m, lst.m (all in d#/proc directory)

%Create list of raw 10 May 1.3 meter frames
loc = '/mnt/mst/1.3meter/y08d131-B'; %location of frames
file = sprintf('%s/chip*', loc); %frame names (e.g. chip_0.032)
fx = dir(file); %Create structure fx with name of each 1.3 meter 10 May frame
l = length(fx); %number of frames
day = 131; %day number
date = [10 05]; %Date of observations

for i = [24,26:75] %Science frames (1-25 on 10 May were test frames)
  n = fx(i).name;  %List of names
  ep = day/365+2008; %epoch to unit day precision

  %extract frame number from frame name
  [a,b,c] = fileparts(n); %separate name into 3 parts (chip, 0, #)
  frame = str2num(c(2:4)); %frame number

  %Read in image to extract times from metadata columns
  im = sprintf('%s/%s', loc, n); %image location and name
  img = fitsread(im); %read image into matlab

  cd d131/proc %Change to proc directory

  %Time data are located either in columns 3 and 4 or 7 and 8
  %Time stamp for each row is formatted in ticks - 0.0002 ticks per second
  %cols 4 or 8 have multiples of 2^16, cols 3 or 7 have 1-2^16
  %See program getrowtime.m for more details

  %determine whether columns 3 and 4 or 7 and 8 have time data
  %the number of ticks per row is 243.86 for 1.3 m clock rate of 48772 microsec/row
  %if the rate of change of column 7 is not 243.86, timing data are in cols 3 and 4
  if median(diff(img(:,3)))>245 || median(diff(img(:,3)))<242
     lsb = img(:,7);
     msb = img(:,8);
  else
       lsb = img(:,3); %read housekeeping column, least significant time bit
       msb = img(:,4); %most significant time bit
  end
```

```matlab
%Call program getrowtime.m to get rowdump time of central row.
time = getrowtime(lsb, msb, 2051);

jd = julday1([date 2008 time/24]); %Call program julday1.m to calculate Julian time

%call program lst.m to get LST of center row, fraction of a day
rl = lst(jd, -1.947787, 'a');
rowf = rl*23.93446959; %fractional hours lst - 23.93446959 sidereal hours per day
lstdeg = rowf*15; %RA degs - 360 degs/24 hours = 15
rowh = floor(rowf); %integer hours
rowm = floor((rowf-rowh)*60); %integer minutes
rows = round(((rowf-rowh)*60-rowm)*60); %integer seconds
rowsf = (((rowf-rowh)*60-rowm)*60); %seconds+frac of sec

cd ../.. %Change back to 1p3m directory

%create RA string in sexagecimal format for new frame name
lstr = sprintf('%02.0f%02.0f%02.0f', rowh, rowm, rows);

%create new name for each frame in format:
%USNO13r2_YYYYDDMMLHHMMSS.fit
%includes observatory (USNO), telescope (13), filter (r), CCD (1.3 m only, 2)
%Date (YYYYMMDD) and RA (HHMMSS).
%L separates date from RA and stands for LST.
newname = sprintf('cp %s/%s d131/proc/USNO13r2_2008%02.0f%02.0fL%s.fit', ...
    loc, n, date(1), date(2), lstr);
system(newname); %calls above unix command to copy the frame to the new name

cd d131/proc

%create RA string in sexagesimal with colons
ra = sprintf('%02.0f:%02.0f:%06.3f', rowh, rowm, rowsf);
%new name
nn = sprintf('USNO13r2_2008%02.0f%02.0fL%s.fit', date(1), date(2), lstr);
%Input RA, Dec, epoch, scale, filter, frame number, time, and RA in degs
%into image header
com = sprintf('sethead %s RA=""%s"" DEC=""35:17:22"" EPOCH=%4.5f
  SECPIX=0.599 FILTER="r" FRAMENUM=%1.0f ROWTIME=%1.12f
  RLSTDEG=%1.9f', nn, ra, ep, frame, time, lstdeg); %setheader
system(com);

cd ../..
fprintf('completed %1.0f of %1.0f files\n', i, l); %Output status of program run.
end
```

Appendix D *biasvec.m*

```
%biasvec.m - creates a bias vector from scan mode bias (dark) frames taken
%at the start of the night

%calls matlab program sigclip.m
%Prior to running copy raw bias frames to d131/bias directory

imgs = dir('chip*'); %Create a list of all bias frames in folder
sim = []; %initialize sim vector
for i = 1:10 %10 bias frames
   im = fitsread(imgs(i).name); %read in bias image

   %Columns in overscan region are a true measure of the instrumental bias
   %Prior to any other processing, remove any changes in bias as a function of row
   %based on changes in the overscan
   over = im(:,[2104:2116]); %overscan region of frame

  %Find the mean counts across the central few columns of overscan region for each row
   m = mean(over(:,7:13),2);
   p = polyfit([1:4102]',m,1); %linear fit to mean overscan counts (overscan vs row)
   v = polyval(p,[1:4102]'); %4102 row vector of overscan trend
   vm = repmat(v, 1, 2116); %copy overscan trend vector to create 4102X2116 matrix
   img = im-vm; %Subtract overscan trend from bias image
   clear im over m p v vm

   sim = [sim; img]; %concatenate bias frames into single matrix (Nx2116)
   clear img
end

clear i imgs

%Trim bias strip image to image pixels (remove over and underscan)
tsim = sim(:,[56:2103]);
clear sim

%sigma clip each column of the bias strip to remove cosmic rays and other anomalies
%and create a 1X2048 bias vector containing a mean bias value for each column.
for i = 1:size(tsim,2)
   clip = sigclip(tsim(:,i)); %sigma clipped vector
   %Find mean of each sigma-clipped column - this is the final bias vector
   bvec(i) = mean(clip);
   len(i) = length(clip); %Length of sigma-clipped column - for debugging only
end

clear i clip
```

Appendix E *flat.m*

```
%flat - Program to create a drift-scan flatfield vector from science frames

%Takes around 5 carefully selected image frames from a single night
%Frames should have similar background counts (no moon, no ramp frames)
%Concatenates bias corrected frames into single strip image then takes
%histograms of each column.  Histogram is dominated by the sky background counts
%which have a parabolic distribution (when log plotted) centered on the mode of
%the background.  A vector of background modes for each column is the flatfield.
%-------------------------------------------------------------------------------%

%Change to directory with raw 1.3 meter 10 May frames and load selected frames
cd /mnt/mst/1.3meter/y08d131-B

sim = []; %initialize strip image matrix
imgs = dir('chip*'); %list of all images in directory
for i = [40,50,60,70,80]; %Selected frames
   im = fitsread(imgs(i).name); %Read in image

   %Correct for overscan region trends as a function of row
   over = im(:,[2104:2116]); %overscan region of frame
   m = mean(over(:,7:13),2); %Mean counts in overscan vs. row
   p = polyfit([1:4102]',m,1); %Fit line to overscan counts vs. row
   v = polyval(p,[1:4102]');
   vm = repmat(v, 1, 2116); %Create matrix with overscan correction
   img = im-vm;  %Apply overscan correction to selected frames

  %Load bias vector and subtract bias counts from images
   load proc/bias131
   bmat = repmat(bvec, 4102, 1); %create bias matrix
   gimg = img(:,[56:2103]); %remove over/underscan
   bimg = gimg - bmat; %subtract bias - bimg is 4102x2048

   clear im bmat gimg img over m p v vm
   sim = [sim; bimg]; %Create concatenated strip image
   clear bimg
end
clear i imgs

for i = 1:2048
   [h, b] = hist(sim(:,i), 200:400); %Histogram of strip centered on sky counts
   a = find(h > 0.1*max(h)); %Find points in histogram within 10% of sky peak

   %Remove points not associated with sky background
   c = find(b(a)<mean(b(a))+3*std(b(a)) & b(a)>mean(b(a))-3*std(b(a)));
```

```matlab
    %fit parabola to log distribution
     p = polyfit(b(a(c)), log(h(a(c))), 2);

     %Create polynomial
     %Vector of values between min and max of sky dist.
     x = [min(b(a(c))):0.01:max(b(a(c)))];
     y = p(1).*x.^2 + p(2).*x + p(3); %Parabola fit to histogram distribution
    %Peak of parabola is where derivative of y crosses 0
     pind = find(diff(y)<0, 1, 'first');
    %Peak of histogram of each column is modal sky background value - flatfield
     peak(i) = x(pind);

     clear h b a c p x y pind

end
clear i

cd /data/flagstaff508/1p3m/d131 %Return to d131 directory
```

Appendix F *seimw.m*

```
%seimw - Program to correct frames for bias, dark and flatfield
%Run source extractor to find positions of stars in frame then run imwcs to find an initial
%coordinate transformation from pixel values to world coordinates (RA, Dec).
%Calls Source Extractor (Bertin and Arnout 1996),
%imwcs (WCSTools package, Mink 1997)
%and Matlab programs fitsheader.m and writecube.m

fs = sprintf('USNO10r_2008*'); %Name of fits files
fx = dir(fs); %Structure containing names of all fits files in directory
for j = 1:length(fx)
    name = fx(j).name; %Image name
    im = fitsread(fx(j).name); %read in image

    %--------------------1.0 meter only--------------------------------%
    %create weight map to flag bad pixels - bad pixels have zero counts and result from an
    % image readout glitch with the 1.0 m CCD controller.  Bad pixels are grouped in lines
    % that are typically 32 pixels long and are aligned with rows.  In current stage the
    % weight map is a structure containing lists of pixel positions corresponding to bad
    % pixels in each frame.  In later program map will be turned into a fits image.

    wmap = []; %Initialize wmap matrix
    for q = 1:4100
        a = find(im(q,17:2048)==0); %Find all pixels with zero counts in row
        if isempty(a)
            clear a
            continue
        else
            for p = 1:length(a) %list bad pixels
                row(p) = q; %Row index of bad pixels
                col(p) = a(p); %column index of bad pixels
            end
        end
        wmap = [wmap; row', col']; %List of positions of all bad pixels in frame
        clear a row col
    end
    clear q

    weight(j).lst = name(18:23); %RA of frame (for later identification)
    weight(j).map = wmap; %List of bad pixels
    clear wmap
    %---------------------------------------------------------------------%

    %Correct for overscan bias variations
    over = im(:,[2065:2112]); %overscan region of frame
```

```
m = mean(over(:,38:45),2); %Take mean across columns of overscan
p = polyfit([1:4100]',m,1); %Fit line to overscan trend vs. row
v = polyval(p,[1:4100]');
vm = repmat(v, 1, 2112); %Create matrix with overscan correction
img = im-vm;  %Apply overscan correction to frame
clear im over m p v vm

gpix = img(:,17:2064); %trim image to science pixels (remove over/underscan)

%load bias and flatfield
load bias135
bmat = repmat(bvec, 4100, 1); %copy bias row to create 4100x2048 matrix
bimg = gpix - bmat; %subtract bias from frame - bimg is 4100x2048
clear img bmat bvec gpix

load flat135
%copy normalized flatfield row to create 4100x2048 matrix
fmat = repmat((peak./mean(peak)), 4100, 1);
fimg = bimg./fmat; %divide flatfield from image
clear fmat peak bimg

%write new fits file with corrected image
%calls program fitsheader, creates stucture with header info
head = fitsheader(fx(j).name, 'fullheader');
%Remove unnecessary comment fields from header
header = rmfield(head, 'COMMENT');
header.BITPIX = -64; %change to 64 bit pixels
header.NAXIS1 = 2048; %change naxis1 value from 2112 to 2048

[junk, file, ext] = fileparts(name); %Extract root of file name
newname = sprintf('%sBF.fits', file(1:23)); %new image name

cd bf %Move to bf directory
writecube(fimg, newname, header) %Write new image
clear head header junk file ext

%rotate resulting file to align with catalog orientation and find world coordinates
c1 = sprintf('imrot -l -r 270 %s', newname); %mirror image left and rotate 270 degs
[stat, rname] = system(c1); %rname is name of rotated image
movefile(rname(1:35), '../sewcs') %move rotated image to sewcs directory

cd ../sewcs %Change to sewcs directory

%Run Source Extractor on frame to find star positions: output is x, y and mag
%This SE run simply provides a list of stars in the frame for imWCS to use which is
% faster than making imWCS find the positions.
```

```matlab
c2 = sprintf('sex -c flag.sex %s', rname(1:35));
[stat, out] = system(c2);

%Run imwcs (WCSTools) on frame:
% -o %sW.fits - write resulting image to root name with W.fits appended
% -c ucac2 - match frame to UCAC2 catalog
% -n 8 - Parameters fit to plane tangent WCS
%        %- center, plate scale (x, y, cross terms), reference coordinates
% -h 200 - Use up to 200 reference stars (max allowed)
% -y 2.0 - Search within area twice the image size
% -q i2sp - iterate fit, sigma clip to 2 stds and fit a quadratic polynomial
% -d test.cat - use source extractor produced list of stellar positions to find
%   %stars in image
% %s - string with name of image to be fit
c4 = sprintf('imwcs -o %sW.fits -c ucac2 -n 8 -h 200 -y 2.0 -q i2sp -d test.cat %s'...
    , fx(j).name(1:23), rname(1:35));
%Output file USNO10r_20081005LHHMMSSW.fits in sewcs directory
[stat, out] = system(c4);

c6 = sprintf('%s', rname(1:35)); %name of rotated image
delete(c6)  %delete rotated image

cd ../bf
c7 = sprintf('%sBF.fits', fx(j).name(1:23));
delete(c7) %delete bias/flatted image
clear stat rname c* out

cd ../sewcs
%get information of quality of fit from header
c8 = sprintf('gethead %sW.fits WCSMATCH', fx(j).name(1:23));
[stat, a] = system(c8);   %Number of reference stars matched
c9 = sprintf('gethead %sW.fits WCSNREF', fx(j).name(1:23));
[stat, b] = system(c9);  %Number of reference stars

%ratio of matches to stars - fitting to wrong stars if very low
effic(j) = str2num(a)/str2num(b);
%Status of program run
c10 = sprintf('completed %1.0f of %1.0f files, efficiency %2.0f/%2.0f = %1.3f for
     %sW.fits', j, length(fx), str2num(a), str2num(b), effic(j), fx(j).name(1:23));
disp(c10)  %Display status of program run

cd ..

clear wname c* stat out name rname a b file junk ext g* newname
end
```

Appendix G *prep13_2.m*

```
%prep13_2.m - Program to create structure containing star positions (pixels) and RA/Dec
of %matched star in catalog as well as errors and other data for each frame
%Calls programs Source Extractor (Bertin and Arnout 1998) and WCSTools program
%imcat (Mink 1997), as well as matlab programs precess2.m, starpix.m and sigclip.m

date = [12 5 2008]; %date of observation (for precession)
%Location and names of processed images
fstr = sprintf('/mnt/mst/1.3meter/processed/d133/USNO*W.fits');
flist = dir(fstr); %List of fits files in directory
l = length(flist); %Number of frames
z = 0; %Initialize index
for i = 1:l
   name = flist(i).name; %Name of frame


   %------------------------------------------------------------------------------------------------------%
   %Run WCSTools program imcat to find UCAC2 catalog stars in image
   % -c ucac2 - catalog called
   % -d - output UCAC2 star positions in decimal degrees
   % -n 5000 - Find and list up to 5000 stars per frame
   % > mlist1.cat - Write output to file mlist1.cat
   %Resulting list has UCAC2 ID, UCAC2 RA, UCAC2 Dec, magj, magh, magk, magc
   %    x-pixel, y-pixel
   %RA and Dec are catalog values of ucac2 stars, x and y are pixel positions of ucac2
   %    stars transformed from RA and Dec using image WCS
   %------------------------------------------------------------------------------------------------------%
   c1 = sprintf('imcat -c ucac2 -d -n 5000 /mnt/mst/1.3meter/processed/d133/%s >
       mlist1.cat', name);
   system(c1);


   %Read resulting mlist1.cat into matlab and format it into a new matrix with positions
   % precessed to observation epoch and which can be read by source extractor.  Also
   % read in u2list.mat which contains a list of all ucac2 stars with full data including
   % errors in matrix u2a

   fid = fopen('mlist1.cat', 'rt'); %read mlist into matlab
   m = textscan(fid, '%f %f %f %*f %*f %*f %f %f %f');
   cm = cell2mat(m); %convert cell to a matrix
   id = cm(:,1); %ucac2 id
   fclose(fid);
   clear fid
   zone = floor(id); %ucac2 zone
   % Determine UCAC2 zone
   for j = 1:length(zone)
```

```matlab
    if zone(j) == 250
       n0(j) = 43899332;
    elseif zone(j) == 251
        n0(j) = 44071847;
  elseif zone(j) == 252
        n0(j) = 44243601;
      end
 end

%convert id to ucac2 standard
stid = round((id-zone).*10^6 + n0');
clear m id zone n0 j

%read in u2list
load u2list

r = cm(:,2); %RA column (deg)
d = cm(:,3); %Dec column (deg)

%Get UT time of center row
c0 = sprintf('gethead ROWTIME /mnt/mst/1.3meter/processed/d133/%s', name);
[stat, t] = system(c0);
time = str2double(t)./24;  %convert to time to a number - fraction of day
tjd = julday1([date, time]);  %Julian date of observation

%find u2list stars in mlist
[c, il, im] = intersect(u2a(:,11), stid);

%-------------------------------------------------------------------------------------------%
%precess RA and Dec to observation epoch using USNO's novas routine
%inputs to mexgetprec are:
%RA in fractional hours
%Dec in fractional degrees
%pmRA in seconds of TIME per CENTURY
%pmDec in arcseconds per CENTURY
%Returns: pRA (frac_hours), pDec (frac_degs)
%calls c-programs mexgetprec.c novas.c novascon.c solsys3.c readeph0.c
%Compile matlab mex C programs using command (in matlab):
%mex -cxx -lm mexgetprec.c novas.c novascon.c solsys3.c readeph0.c
%Applies precession to observation epoch as well as proper motion
%-------------------------------------------------------------------------------------------%
cd /data/flagstaff508/novas-c201
for j = 1:length(im);
   [pr(j), pd(j)] = mexgetprec(tjd, r(im(j))*24/360, d(im(j)),
      u2a(il(j),7)*36./(15./cos(d(im(j))*pi/180)), u2a(il(j),8)*36);
end
```

```
clear j
cd /data/flagstaff508/1p3m/d133/proc/Gfit

%convert ucac2 ID (8 digits) to two numbers
id = num2str(stid(im)); %convert to string
id1 = str2num(id(:,1:4)); %first 4 digits of ID
id2 = str2num(id(:,5:8)); %second 4 digits of ID

fid = fopen('mlist.cat', 'wt'); %Open a new text file mlist.cat for writing
for j = 1:length(c);
%Write id1, id2, precessed RA, precessed Dec, mag, x, y, RA_error, Dec_error, pmra,
%    pmdec, pmra_err, pmdec_err, cera, cedec to file
    fprintf(fid,'%4.0f\t%4.0f\t%1.8f\t%1.8f\t%f\t%1.2f\t%1.2f\t%1.8f\t%1.8f\t%1.5e\
            t%1.5e\t%1.5e\t%1.5e\t%4.4f\t%4.4f\n',  id1(j), id2(j), pr(j).*360/24,
            pd(j), cm(im(j),[4,5,6]), u2a(il(j),[3,4,7:10,5,6]));
end
fclose(fid);

%Same matrix of imcat data (as above)
mlist = [u2a(il,11), pr'.*360/24, pd', cm(im,[4,5,6]), u2a(il,[3,4,7:10,5,6])];
clear cm stid u2a fid j c1 r d pr pd t c0 stat il im c


%-------------------------------------------------------------------------------------------------%
%Run Source extractor on frames using the ASSOC feature to match stars in mlist.cat
%with stars in the image and write the result to test.cat
%Resulting text file has columns: XWIN YWIN Flux_auto Fluxerr_auto Flags
%    Mag_auto Back Thresh X2_image Y2_image X2err Y2err FWHM_image
%    ucac2_id1, id2
%X/YWIN - positions are calculated within a circular gaussian aperture
%AUTO - flux and magnitudes are calculated within an ellipse determined by the
%    object's second order image moments
%Flags indicate possibly compromised (saturated, blended, etc.) objects.  0 good.
%Back and Thresh are sky background and detection threshold above sky background
%X2/Y2_image are second order image moments (pixels)
%FWHM is full width at half maximum of object (pixels)
%ucac2_id is ucac2 id of star associated with source extractor obj
%-------------------------------------------------------------------------------------------------%
c0 = sprintf('sex -c flag.sex /mnt/mst/1.3meter/processed/d133/%s', flist(i).name);
[stat,out] = system(c0); %Source extractor output

fid = fopen('test.cat', 'rt'); %read SE list into matlab
%Read in list with tab delimiters and 14 headerlines (headerlines are discarded)
c = textscan(fid, '%f %f %f %f %f %f %f %f %f %f %f %f %f %f %f', ...
        'delimiter', '', 'headerLines', 14);
fclose(fid);
cat = cell2mat(c); %convert resulting cell structure to a matrix
```

```
clear fid c

c = find(cat(:,5) == 0); %Find all stars with 0 flags (non-compromised objects)
gcat = cat(c,[1:4,6:15]); %remove SE flagged stars
%gcat contains: x y flux fluxerr mag back thresh x2 y2 xerr yerr fwhm ucac2id

im = sprintf('/mnt/mst/1.3meter/processed/d133/%s', name);  %Name of image
img = fitsread(im); %Read image into matlab

p = 0; %initialize p index
%Run program starpix on all stars to remove saturated or low s/n objects
for k = 1:size(gcat,1)
    %input to starpix is image, x, y, flux, fluxerr
         [good, sn, peak] = starpix(img, gcat(k,1), gcat(k,2), gcat(k,3), gcat(k,4));
         if good == 0 %if good = 0, star is saturated or low s/n, discard
       clear good sn peak
       continue
    end
    p = p+1; %increment index for each acceptable star

    %Create new matrix with acceptible stars, columns: X Y flux mag
    %int s/n back x2 y2 x_error y_error fwhm ucac2id1/2
    bcat(p,:) = [gcat(k,[1,2,3,5]), peak, sn, gcat(k,[6,8:14])];
    clear good peak sn
end

clear c im* p k gcat

%Call program sigclip.m to remove objects with anomalous flux/intensity peak ratios -
%     possibly galaxies or blended objects
clip = sigclip(bcat(:,3)./bcat(:,5));
[c, is, ib] = intersect(clip, bcat(:,3)./bcat(:,5));
bcat = bcat(ib,:);

idb = bcat(:,13).*10^4+bcat(:,14); %convert two part ucac2id back to single number
%Find all objects in imcat list, mlist, that match objects in SE list, bcat
[id, sc, nc] = intersect(idb, mlist(:,1));
clear c is ib clip

%Create new list of resulting objects in image and associated catalog objects
%Columns are: ID X Y RA Dec mag X_err Y_err RA_err Dec_err fwhm x2 y2 flux
%     s/n back pmra pmdec epmra epmdec cera cedec
%X, Y, X/Y_err, fwhm, x/y2, flux, back and s/n are from Source Extractor list
%RA, Dec, RA/Dec_err, mag, proper motions and errors, and central epochs are from
%    imcat list
table = [id, bcat(sc,1:2), mlist(nc,2:4), bcat(sc,10:11), mlist(nc,7:8), ...
```

```matlab
    bcat(sc,12), bcat(sc,8:9), bcat(sc,[3,6,7]), mlist(nc,9:14)];

    %Get central RA and Dec of frame from header
    com = sprintf('gethead /mnt/mst/1.3meter/processed/d133/%s XCPOS', name);
    [stat, out] = system(com);
    com2 = sprintf('gethead /mnt/mst/1.3meter/processed/d133/%s YCPOS', name);
    [stat, out2] = system(com2);
    z = z+1; %Increment index for usable frames (more than 10 stars)

    %Structure with a list of stars, their info, central RA/Dec and name for each frame
    frame(z).cat = table;
    frame(z).cRA = str2double(out);
    frame(z).cDec = str2double(out2);
    frame(z).name = name;

    %Precess central frame RA/Dec to observation epoch
    cd /data/flagstaff508/novas-c201
    [pRA, frame(z).pDec] = mexgetprec(tjd, frame(z).cRA*24/360, frame(z).cDec, 0, 0);
    frame(z).pRA = pRA*360/24; %convert central precessed RA to deg.
    cd /data/flagstaff508/1p3m/d133/proc/Gfit

    %Print status of program run - completed frame number and number of stars in frame
    fprintf('completed %1.0f of %1.0f frames, %1.0f stars\n', i, l, size(table,1))
    clear out* idb stat c* table a q j b slist bcat ind outmat id sc pRA tjd nc mlist name
time
end

clear z i flist l fname fstr ans epoch
```

Appendix H *mbmaster.m*

```
%mbmaster - universal m-file for taking calibration and mb data and processing it.
clear; help mbmaster

%-------------------------------------------------------------------------------------------------------%
%Calibration data includes pressure data obtained  with no pressure input (for zero
%offset), and data
%obtained with several different pressure inputs to determine any variation in pressure
%response.
%Temperature calibration data will be obtained and processed separately on occasion to
%calculate temp error functions
%Filenames:
%       cal - pressure calibration data
%       pistpos - piston position (input pressure) during calibration program
%       temp
%       weather - weather data
%       pressure - field pressure data
%-------------------------------------------------------------------------------------------------------%

% Calibration data
cname = 'cal';
cfile = sprintf('%s*', cname);
cx = dir(cfile);
fid = fopen(cx.name);
%Read in calibration data text file: columns are time, p1, p2, p3, p4
cc = textscan(fid, '%f %f %f %f %f');
fclose(fid);
%parse filename for starttime of calibration
[junk, f, ext] = fileparts(cx.name);
[junk, sttime] = strtok(f, '_');
cstarttime = datenum(sttime, '_yy-mm-dd_HHMM');
cal = cell2mat(cc);
%Convert time vector (seconds) to MATLAB datenum format
ct = cstarttime + (cal(:,1).*1.157412771135569e-005);

%Convert pressure voltage readings to millibars
%-----------------------------------------%
%Config F: p = 0.498*(volts - 2.5)
%Config G: p = 0.249*(volts - 2.5)
%Config H: p = 0.1245*(volts - 2.5)
%-----------------------------------------%
cp4 = (cal(:,5)-2.5).*0.498; %DPT5 (millibars)
cp1 = (cal(:,2)-2.5).*0.1245; %DPT1
cp2 = (cal(:,3)-2.5).*0.1245; %DPT3
cp3 = (cal(:,4)-2.5).*0.1245; %DPT4
```

```
clear junk f ext sttime fid cc cx cfile cname cal

%read in temperature data
Temp = [];
Tt = [];
tname = 'temp_08';
tfile = sprintf('%s*', tname);
tx = dir(tfile);
for i = 1:length(tx)
    fid = fopen(tx(i).name);
 %read in temperature data text file: columns are time, temp.
    tc = textscan(fid, '%f %f');
    fclose(fid);
    temp = cell2mat(tc);
 %parse filename for starttime of data
    [junk, f, ext] = fileparts(tx(i).name);
    [junk, sttime] = strtok(f, '_');
    tstarttime = datenum(sttime, '_yy-mm-dd_HHMM');
        %Convert time vector (seconds) to MATLAB datenum format
    Tt = [Tt; tstarttime + (temp(:,1).*1.157412771135569e-005)];
        %Convert temperature voltages to degrees celsius
    Temp = [Temp; (temp(:,2)- 0.805858)./(-0.0056846)];
    clear fid tc temp junk f ext sttime tstarttime
end
clear i tname tfile tx

%determine temperature data point corresponding to start of calibration sequence
st = find(Tt < ct(1), 1, 'last');
if isempty(st)
    st = 1;
    cs = find(ct < Tt(1), 1, 'last');
    ct = ct(cs:length(ct));
end
%determine temperature data point corresponding to end of calibration sequence
en = find(Tt > ct(length(ct)), 1, 'first');
%Interpolate temperature points to the same grid as the calibration pressure points
cTemp = interp1(Tt(st:en), Temp(st:en), ct);

clear st en temp tc fid tx tfile tname junk f ext sttime tstartime

% Correct for temperature error based on calculated temperature dependeces of the DPTs
cp1T = cp1 - (0.00034*cTemp - 0.002); %DPT#1
cp3T = cp3 - (0.00058*cTemp - 0.023); %DPT#3
cp2T = cp2 - (0.0037*cTemp - 0.17); %DPT#4 - new style, approx
cp4T = cp4;
```

```
%in workspace:  cstarttime, ct, cp1T, cp2T, cp3T, cp4T, cTemp, Tt, Temp

% Read in weather data
name1 = 'weather';
wfiles = sprintf('%s*', name1);
wx = dir(wfiles);

date = []; %Date in format YYYYMMDD-hhmmss
patm = []; %Atmospheric pressure in Hg
tout = []; %Atm. temp in deg. C
wind = []; %Wind speed in mph
dire = []; %Wind direction in deg
rh = []; %Relative humidity: perc.

%Weather files start at 00:00 - 2 files for each night
for i = 1:length(wx)
   fid = fopen(wx(i).name);
  %read in weather data: columns are date/time, pressure, temp, wind speed, wind
  %direction, relative humidity
   cw = textscan(fid, '%*s %f %f %f %f %f');
   fclose(fid);
   dt = textread(wx(i).name, '%s %*f %*f %*f %*f %*f');
   w = cell2mat(cw);
 %convert date and time to matlab dates and times
   date = [date; datenum(dt, 'yyyymmdd-HHMMSS')];
   patm = [patm; (w(:,1).*3.3769e3)]; %Atmospheric pressure (Pa)
   tout = [tout; (w(:,2))]; %Outside temp (deg. C)
   wind = [wind; w(:,3)]; %wind speed (mph)
   dire = [dire; w(:,4)]; %wind direction (deg)
   rh = [rh; (w(:,5)./100)]; %Relative humidity (frac)
end

clear n* wf* wx i f* cw dt w
%in workspace:  cstarttime, ct, cp1T, cp2T, cp3T, cT, date, patm, rh, wind, dire

%remove any DC offsets
cp1Tm = cp1T - mean(cp1T);
cp2Tm = cp2T - mean(cp2T);
cp3Tm = cp3T - mean(cp3T);
cp4Tm = cp4T - mean(cp4T);

clear b cp1T cp2T cp3T cp4T a ic ip
%in workspace: cstarttime, ct, cp1Tm, cp2Tm, cp3Tm, cp4Tm, Tt, Temp

%Find vector of mean DPT response
mcal = mean([cp1Tm cp2Tm cp3Tm], 2);
```

```matlab
%create linear fit to response data vs. mean (cp1Tm = m*mcal+b)
pf1 = polyfit(mcal, cp1Tm, 1); %pfn = [m, b];
pf2 = polyfit(mcal, cp2Tm, 1);
pf3 = polyfit(mcal, cp3Tm, 1);
pf4 = polyfit(mcal, cp4Tm, 1);

%To correct data: pcorr = (p - pf(2))/pf(1)

p1 = [];
p2 = [];
p3 = [];
p4 = [];
time = [];

%Read in field data
name2 = 'pressure';
pfiles = sprintf('%s*', name2);
px = dir(pfiles);
for i = 1:length(px)
   fid = fopen(px(i).name);
  %read in text file of pressure data: columns are time p1 p2 p3 p4
   cp = textscan(fid, '%f %f %f %f %f');
   fclose(fid);
   p = cell2mat(cp);
   ti = p(:,1);
   t = ti - ti(1);
  %parse file name for start time
   [junk, f, ext] = fileparts(px(i).name);
   [junk, sttime] = strtok(f, '_');
   starttime = datenum(sttime, '_yy-mm-dd_HHMM');
        %convert time vector (seconds) to matlab datenum format
   sec = t.*1.157412771135569e-005;
   pt = starttime + sec;  %time of pressure reading
   time = [time; pt];
        %convert pressure voltages to pressures
   p1 = [p1; (p(:,2)-2.5).*0.1245];
   p2 = [p2; (p(:,3)-2.5).*0.1245]; %configuration H
   p3 = [p3; (p(:,4)-2.5).*0.1245];
   p4 = [p4; (p(:,5)-2.5).*0.498]; %config F
%------------------------------------------%
%Config F: p = 0.498*(volts - 2.5)
%Config G: p = 0.249*(volts - 2.5)
%Config H: p = 0.1245*(volts - 2.5)
%------------------------------------------%

   clear p pt sec starttime sttime junk f ext t ti fid cp
```

```
end
clear name2 pfiles px i
%find temp data corresponding to pressure data
st = find(Tt <= time(1), 1, 'last');
en = find(Tt <= time(length(time)), 1, 'last');
temp = Temp(st:en);
Ttime = Tt(st:en);

clear name2 P_in pfiles px p pt sec st* fid t j* f ext sttime xp ps* V Mv rh patm mcal en
        Temp Tt
%in workspace:  date wind dire dp* fdp* time p* temp s cp*Tm ct cT pf*

% Remove any bad weather data
%Good wind data has values from 0 to max wind speed (<30 mph because otherwise we
%don't observe)
good = find(wind > -1 & wind < 30); %Careful here if it's really windy!
gtime = date(good);
gwind = wind(good);
gdir = dire(good);
gtemp = tout(good);

% Crop weather and/or pressure vectors to the same time range
startp = 1;
startg = 1;
startT = 1;
if time(1) > gtime(1)
   startg = find(gtime >= time(1), 1, 'first');
elseif time(1) < gtime(1)
   startp = find(time >= gtime(1), 1, 'first');
    startT = find(Ttime >= gtime(1), 1, 'first');
end

lastg = length(gtime);
lastp = length(time);
lastT = length(Ttime);
if time(length(time)) < gtime(length(gtime))
   lastg = find(gtime <= time(length(time)), 1, 'last');
elseif time(length(time)) > gtime(length(gtime))
   lastp = find(time <= gtime(length(gtime)), 1, 'last');
   lastT = find(Ttime <= gtime(length(gtime)), 1, 'last');
end

gtimenew = gtime(startg:lastg);
gwindnew = gwind(startg:lastg);
gdirnew = gdir(startg:lastg);
gtempnew = gtemp(startg:lastg);
```

```
timenew = time(startp:lastp);
Ttimenew = Ttime(startT:lastT);
p1new = p1(startp:lastp);
p2new = p2(startp:lastp);
p3new = p3(startp:lastp);
p4new = p4(startp:lastp);
Tnew = temp(startT:lastT);

clear good st* la* wind dir date p1 p2 p3 p4 temp gtime gwind gdir time Ttime
%in workspace: s gtimenew gwindnew gdirnew timenew p*new Tnew
%cp*Tm ct cT pf*

badst = [];
badend = [];
badTst = [];
badTend = [];
badpst = [];
badpend = [];

x = 0;
%locate gaps in weather data, this is less of an issue in more recent data
for j = 2:length(gtimenew)
  %if more than 10 seconds between two points note start and end of gap
  if gtimenew(j) - gtimenew(j-1) > 10*1.157412771135569e-005
    x = x + 1;
    gapa(x) = gtimenew(j-1); %Time at start of weather data gap
    gapb(x) = gtimenew(j); %Time at end of weather data gap
    %Pressure index corresponding to start of gap
    bad1 = find(timenew >= gapa(x), 1, 'first');
            %if no pressure values at start of gap or later, set badst to -1
    if isempty(bad1)
      badst(x) = -1;
    else
      badst(x) = bad1(1);
      badTst(x) = find(Ttimenew >= gapa(x), 1, 'first');
    end
    %Pressure index corresponding to end of gap
    bad2 = find(timenew >= gapb(x), 1, 'first');
    if isempty(bad2) %if no pressure values at end of gap set badend to -1
      badend(x) = -1;
    else
      badend(x) = bad2(1);
      badTend(x) = find(Ttimenew >= gapb(x), 1, 'first');
    end
  end
end
```

```matlab
clear j bad1 bad2
%In workspace:  dp* fdp* s gtimenew gwindnew gdirnew timenew p*new Tnew gapa
%gapb badst badend cpT* ct P_in

y = 0;
%locate gaps in pressure data
for k = 2:length(timenew)
%if time between two pressure datapoints is > one second, note start and end of gap
   if timenew(k) - timenew(k-1) > 1.157412771135569e-005
      y = y+1;
      pgapa(y) = timenew(k-1); %time at start of pressure gap
      pgapb(y) = timenew(k); %time at end of pressure gap
      %Weather index corresponding to start of gap
      badp2 = find(gtimenew >= pgapa(y));
      if isempty(badp2)
         badpst(y) = -1;
      else
         badpst(y) = badp2(1);
      end
      %Weather index corresponding to end of gap
      badp3 = find(gtimenew >= pgapb(y));
      if isempty(badp3)
         badpend(y) = -1;
      else
         badpend(y) = badp3(1);
      end
   end
end

clear k badp2 badp3
% In workspace:  dp* fdp* s gtimenew gwindnew gdirnew timenew p*new Tnew gapa
% gapb badst badend pgapa pgapb badpst badpend x y cpT* ct P_in

%Mind the gap
lp = length(timenew);
lT = length(Ttimenew);
lw = length(gtimenew);
ptime = timenew;
pp1 = p1new;
pp2 = p2new;
pp3 = p3new;
pp4 = p4new;

clear timenew p1new p2new p3new p4new
%In workspace:  dp* fdp* s gtimenew gwindnew gdirnew Tnew gapa gapb badst
```

253

```
% badend pgapa pgapb badpst badpend lp lw ptime pp* cpT* ct P_in

%Remove pressure data corresponding to gaps in weather data
if isempty(badst) == 1
    str = 'no gap';
else
    for i = 1:x
        if badst(i) == -1 %pressure data missing at start of wx gap, but data at end of gap
            pindst = find(ptime < gapa(i));
            Tst = find(Ttime < gapa(i), 1, 'last');
            bst = pindst(length(pindst));
            ptime = [ptime(1:bst); ptime(badend(i):lp)];
            pp1 = [pp1(1:bst); pp1(badend(i):lp)];
            pp2 = [pp2(1:bst); pp2(badend(i):lp)];
            pp3 = [pp3(1:bst); pp3(badend(i):lp)];
            pp4 = [pp4(1:bst); pp4(badend(i):lp)];
            Tnew = [Tnew(1:Tst); Tnew(badTend(i):lT)];
            Ttimenew = [Ttimenew(1:Tst); Ttimenew(badTend(i):lT)];
        elseif badend(i) == -1 %pressure data at start of gap but missing at end
            pindend = find(ptime > gapb(i));
            Tindend = find(Ttimenew > gapb(i), 1, 'first');
            ptime = [ptime(1:badst(i)); ptime(pindend(1):lp)];
            pp1 = [pp1(1:badst(i)); pp1(pindend(1):lp)];
            pp2 = [pp2(1:badst(i)); pp2(pindend(1):lp)];
            pp3 = [pp3(1:badst(i)); pp3(pindend(1):lp)];
            pp4 = [pp4(1:badst(i)); pp4(pindend(1):lp)];
            Tnew = [Tnew(1:badTst(i)); Tnew(Tindend(1):lT)];
            Ttimenew = [Ttimenew(1:badTst(i)); Ttimenew(Tindend(1):lT)];
        elseif badst(i) == -1 && badend(i) == -1 %No pressure data in gap
            continue
        else %pressure data continuous across wx gap
            ptime = [ptime(1:badst(i)); ptime(badend(i):lp)];
            pp1 = [pp1(1:badst(i)); pp1(badend(i):lp)];
            pp2 = [pp2(1:badst(i)); pp2(badend(i):lp)];
            pp3 = [pp3(1:badst(i)); pp3(badend(i):lp)];
            pp4 = [pp4(1:badst(i)); pp4(badend(i):lp)];
            Tnew = [Tnew(1:badTst(i)); Tnew(badTend(i):lT)];
            Ttimenew = [Ttimenew(1:badTst(i)); Ttimenew(badTend(i):lT)];
        end
        lp = length(ptime);
        lT = length(Ttimenew);
    end
end

clear str i x badst pindst bst lp badend gapa gapb pindend badT* Tindend Tst lT
% In workspace: dp* fdp* s gtimenew gwindnew gdirnew Tnew pgapa pgapb
```

```
% badpst badpend lw ptime pp* y cpT* ct P_in

%remove weather data corresponding to gaps in pressure data
wtime = gtimenew;
wind = gwindnew;
wdir = gdirnew;
wtemp = gtempnew;
if isempty(badpst) == 1
   wstr = 'no gap';
else
   for i = 1:y
      if badpst(y) == -1 %wx data missing at start of press. gap, present at end of gap
         weindst = find(wtime < pgapa(y));
         wbst = weindst(length(weindst));
         wtime = [wtime(1:wbst), wtime(badpend(i):lw)];
         wind = [wind(1:wbst), wind(badpend(i):lw)];
         wdir = [wdir(1:wbst), wdir(badpend(i):lw)];
         wtemp = [wtemp(1:wbst), wtemp(badpend(i):lw)];
      elseif badpend(y) == -1 %wx data missing at end of gap, present at start
         weindend = find(wtime > pgapb(y));
         wtime = [wtime(1:badpst(i)), wtime(weindend(1):lw)];
         wind = [wind(1:badpst(i)), wind(weindend(1):lw)];
         wdir = [wdir(1:badpst(i)), wdir(weindend(1):lw)];
         wtemp = [wtemp(1:badpst(i)), wtemp(weindend(1):lw)];
      elseif badpst(y) == -1 && badpend(y) == -1 %no wx data in pressure gap
         continue
      else %wx data continuous across gap
         wtime = [wtime(1:badpst(i)); wtime(badpend(i):lw)];
         wind = [wind(1:badpst(i)); wind(badpend(i):lw)];
         wdir = [wdir(1:badpst(i)); wdir(badpend(i):lw)];
         wtemp = [wtemp(1:badpst(i)); wtemp(badpend(i):lw)];
      end
      lw = length(wtime);
   end
end

clear g* wstr i y badpst badpend lw wbst weindst weindend pgapa pgapb
%In workspace:  dp* fdp* s wtime wind dir Tnew ptime pp* cpT* ct P_in

%Find temperature datapoints corresponding to science pressure data
a = find(Ttimenew >= ptime(1) & Ttimenew <= ptime(length(ptime)));
b = find(ptime >= Ttimenew(1) & ptime < Ttimenew(length(Ttimenew)));
Ttimenew = Ttimenew(a);

ptime = ptime(b(1:(length(b)-1)));
```

```
%correct data for DPT temp dependence
p1T = pp1(b(1:(length(b)-1))) - (0.00034*Tnew(a) - 0.002); %DPT#1
p2T = pp2(b(1:(length(b)-1))) - (0.00014*Tnew(a) - 0.01); %DPT#2
p3T = pp3(b(1:(length(b)-1))) - (0.00058*Tnew(a) - 0.023); %DPT#3
p4T = pp4(b(1:(length(b)-1))) - (0.0037*Tnew(a) - 0.17); %DPT#4 - new style approx

%Remove any overall offsets
p1Tm = p1T - mean(p1T);
p2Tm = p2T - mean(p2T);
p3Tm = p3T - mean(p3T);
p4Tm = p4T - mean(p4T);
Tnew = Tnew(a);
clear a b cp

% Apply calibration data
p1c = (p1Tm - pf1(2))./pf1(1);
p2c = (p2Tm - pf2(2))./pf2(1);
p3c = (p3Tm - pf3(2))./pf3(1);
p4c =  p4Tm; %(p4Tm - pf4(2))./pf4(1); %p4 is point source config, no calibration data.
```

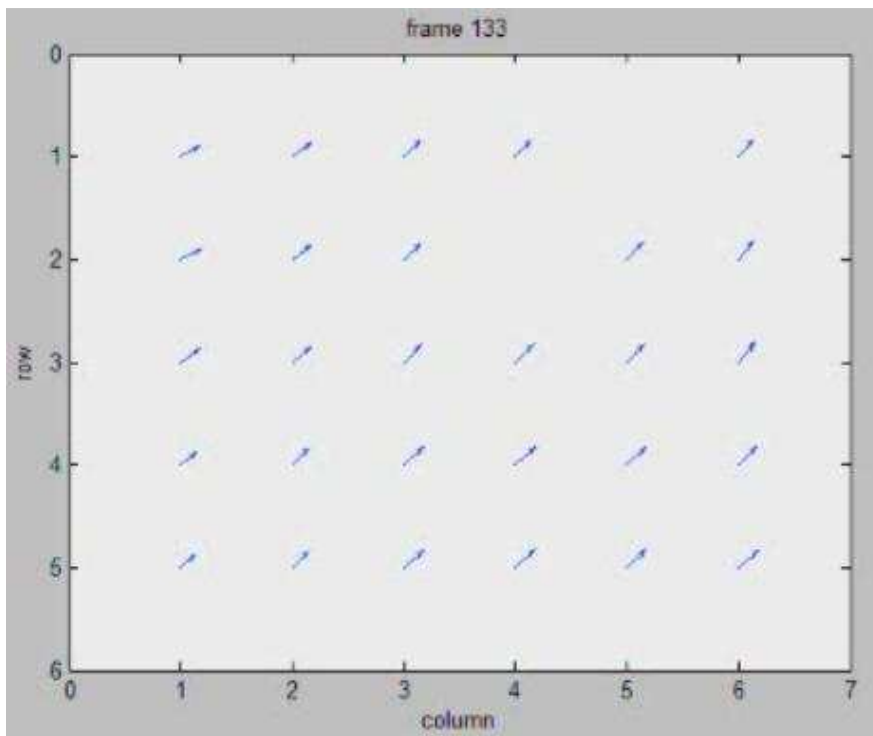# Appendix I *SDSS focal plane map movie images*

Figure 130. Selected frame from the SDSS stripe 82, run 4203 residual focal plane map movie (see Section 3.2.2). Arrows represent the vector sum of RA and Dec residuals for each CCD in the focal plane.
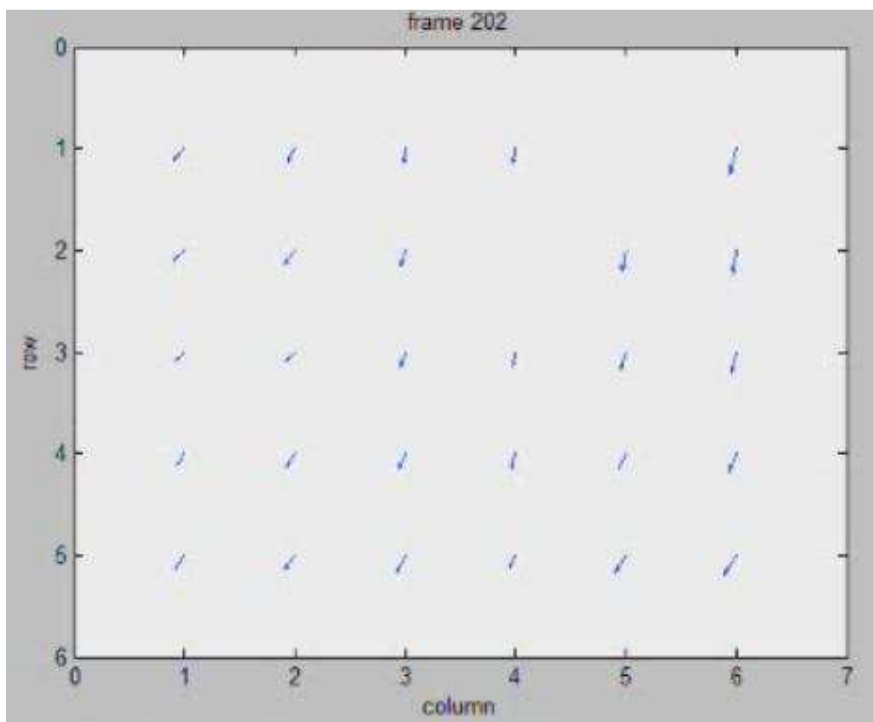

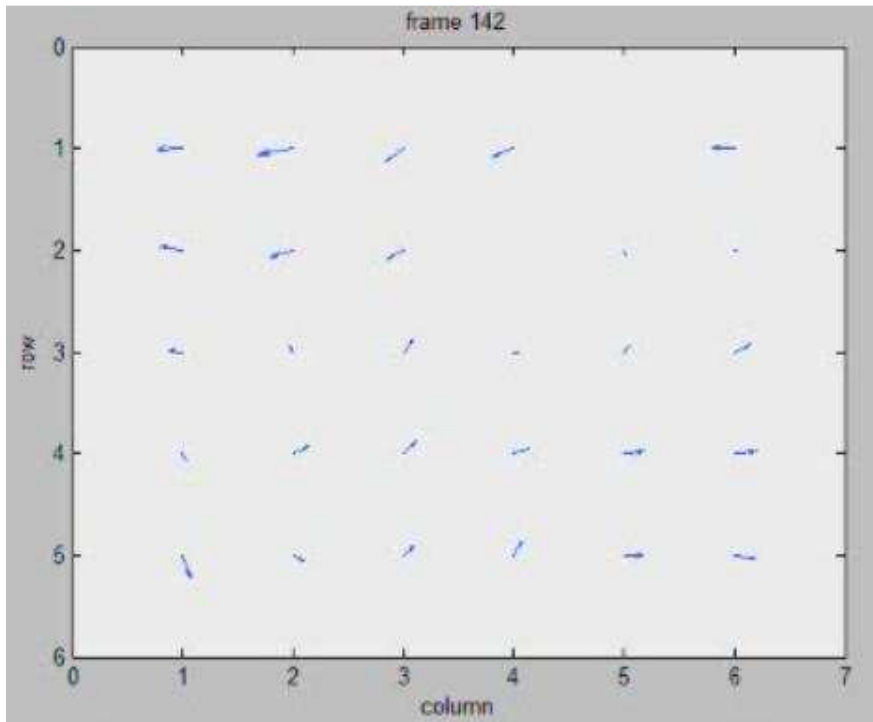**Figure 131.** Later frame from the SDSS stripe 82, run 4203 residual focal plane map movie.

257

Figure 132. Same as above two figures with mean residual (across all CCDs) in each frame subtracted.
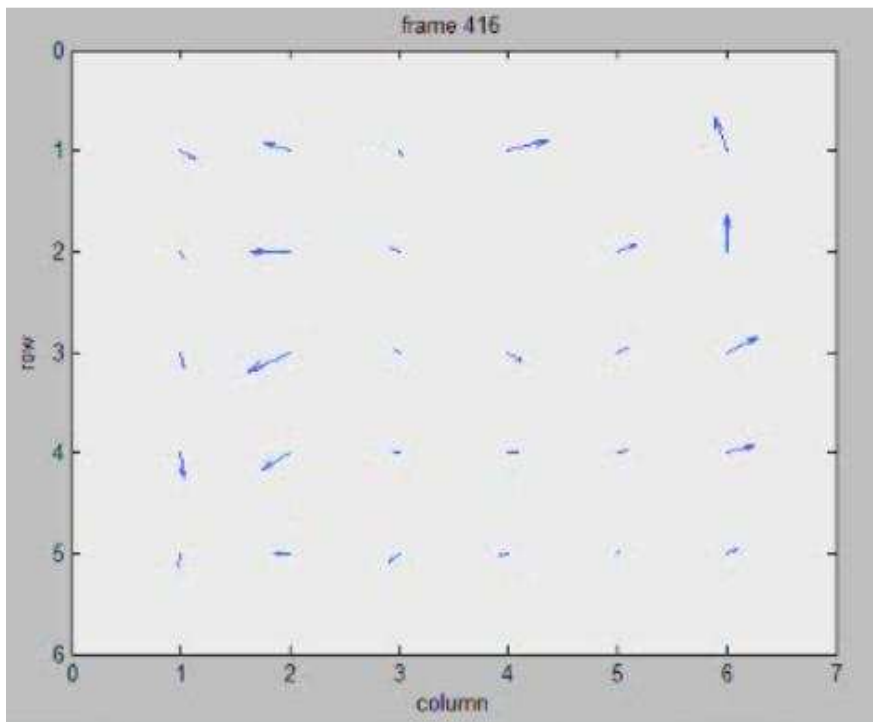

Figure 133. Later frame from same mean-subtracted residual movie as above figure.
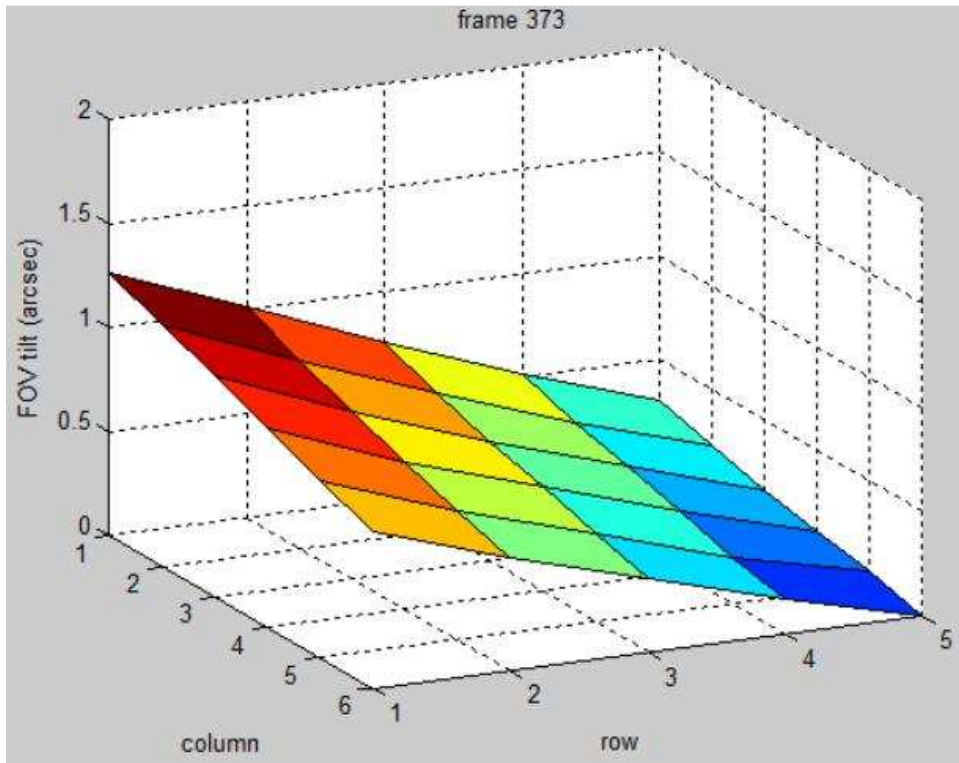
Figure 134. Frame from the SDSS stripe 82, run 5646 focal plane tilt movie (see Section 3.2.2). Surface is the gradient of the residual vectors, representing the apparent tilt of the focal plane.
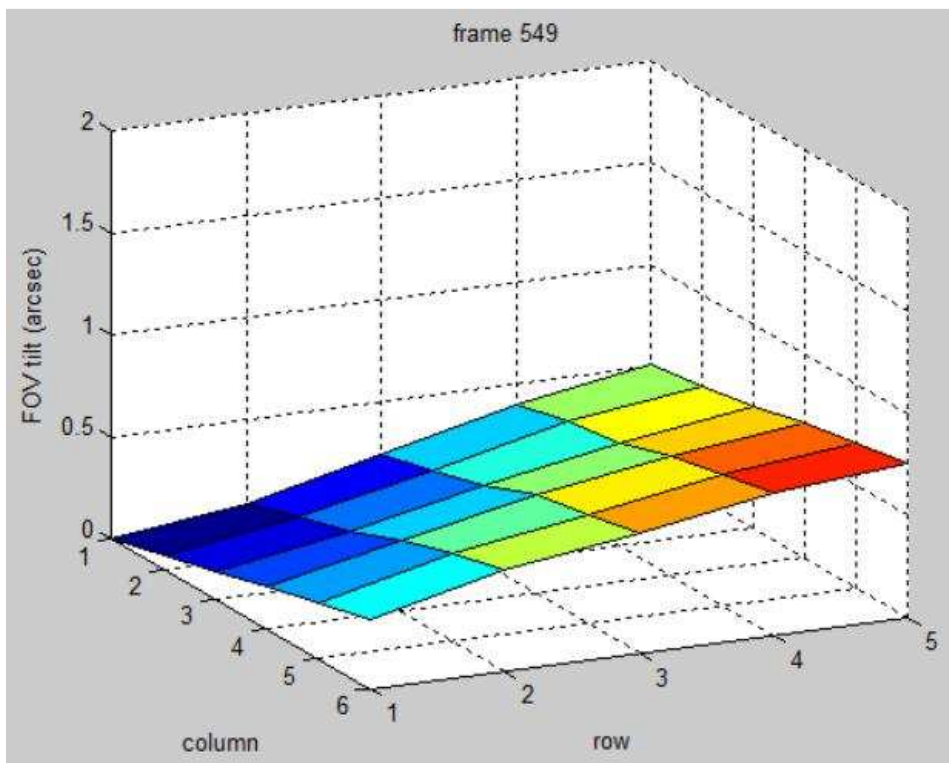

Figure 135. Later frame from the same focal plane tilt movie as the above figure.

259

References

Alcoverro, B. & Le Pichon, A. 2005, J. Acoust. Soc. Am., 117, 1717

Anderson, P.S., Mobbs, S. D., King, J. C., McConnell, I. & Rees, J. M. 1992, Antarctic
Science, 4, 241

Antonia, R. A., Chambers, A. J., Friehe, C. A., & Van Atta, C. W. 1979, J. Atm. Sci, 36,
99

Archimedes, circa 250 BCE, *On Floating Bodies I*

Banta, R. M., Darby, L. S., Newsom, R. K., Hardesty, R. M., & Howell, J. N. 2000, 20th
International Laser Radar Conference, Vichy, France, 10-14 July

Becker, A. C., Silvestri, N. M., Owen, R. E., Ivezic, Z., & Lupton, R. H. 2007, PASP,
119, 1462

Benedict, G. F., McGraw, J. T., Hess, T. R., Cawson, M. G. M., & Keane, M. J. 1991,
AJ, 101, 1

Berger, J. & Levine, J. 1974, J. Geophys. Res., 79, 1210

Bertin, E. & Arnouts, S. 1996, 117, 393

Birch, K. P. & Downs, M. J. 1994, Metrologia, 31, 315

Blumen, W., Banta, R., Burns, S. P., Fritts, D. C., Newsom, R., Poulos, G. S., & Sun, J.
2001, Dynamics of Atm. & Oceans, 34, 184

Bowell, E. L. G. 2005, ASPC, 338, 147

Brunet, Y. & Irvine, M. R. 2000, Boundary-Layer Met., 94, 139

Cava, D., Giostra, U., Siqueira, M., & Katul, G. 2004, Boundary-Layer Met., 112, 129

Chandler, S. C. 1885, Astronomische Nachrichten, 2672, 112

Chandler, S. C. 1887, Ann. Ast. Obs. Harvard, 17, 1

Chandler, K. C. 2004, ASP, 338, 134

Chimonas, G. 1999, Boundary-Layer Met., 90, 397

Ciddor, P. E. 1996, Applied Optics, 35, 1566

Copenhagen Univ. Obs., Inst. of Astronomy, Cambridge, UK, & Real Instituto Y

      Observatorio de La Armada, F. E. S. 2006, VizieR Online Data Catalog, 1304, 0

Cox, A. N. Ed. 2000, Allen's Astrophysical Quantities (New York: AIP)

Cuxart, J., *et al.* 2000, Boundary-Layer Met., 96, 337

Danaila, I., Joly, P., Kaber, S. M., & Postel, M. 2005, An Introduction to Scientific

      Computing (Springer)

Daniels, F. B. 1959, J. Acoust. Soc. Am., 31, 529

Demirgian, J. & Dedecker, R. 2005, Atmospheric Emitted Radiance Interferometer

      Handbook, ARM TR-054

Dunham, E. W., McDonald, S. W., & Elliot, J. L. 1991, AJ, 102, 4

Eaton, F. D., McLaughlin, S. A., & Hines, J. R. 1995, Radio Science, 30, 1

Einaudi, F. & Finnigan, J. J. 1993, 50, 1841

Evans, D. W., Irwin, M. J., & Helmer, L. 2002, A&A, 395, 347

Feltz, W. F., Smith, W. L., Howell, H. B., Knuteson, R. O., Woolf, H., & Revercomb, H.

      E. 2003, J. App. Met., 42, 584

Fried, D. L. 1965, JOSA, 55, 1427

Fritts, D. C. & Alexander, M. J. 2003, Rev. Geophys., 41, 1

Fritts, D. C., Vadas, S. L., Wan, K., & Werne, J. A. 2006, J. Atm. & Sol.-Terr. Phys., 68,

      247

Gehrels, T., Marsden, B. G., McMillan, R. S., & Scotti, J. V. 1986, AJ, 91, 1242

Gibson, B. K. & Hickson, P. 1992, MNRAS, 258, 543

Gossard, E. E. & Munk, W. 1954, Journal of Meteorology, 11, 259

Gossard, E. E. & Sweezy, W. B. 1974, J. Atm. Sci., 31, 1540

Goullioud, R., Dekens, F. G., Nicaise, F., & Nemati, B. 2009, AAS Meeting #213,
     #300.02

Greenwood, D. P. & Fried, D. L. 1976, JOSA, 66, 193

Gubler, J. & Tytler, D. 1998, PASP, 110, 738

Gunn, J. E. *et al.* 1998, AJ, 116, 3040

Guseva, I. S. 1995, IAU, 166, 371

Hardy, J. W. 1998, Adaptive Optics for Astronomical Telescopes (New York: Oxford
     University Press)

Haurwitz, B. 1941, Dynamic Meteorology (New York: McGraw-Hill Book Company,
     Inc.)

Hedlin, M. A. H. & Alcoverro, B. 2005, J. Acoust. Soc. Am., 117, 1880

Hedlin, M. A. H., Berger, J., & Zumberge, M. 2000, Evaluation of Infrasonic Spatial
     Filters, in 22nd Annual DoD/DOE Seismic Research Symposium

Hedlin, M. A. H. & Raspet, R. 2003, J. Acoust. Soc. Am., 114, 1379

Helmert, F. R. & Albrecht, T. 1898, Astronomische Nachrichten, 148, 49

Hennessy, G., Hindsley, B., Kent, S., Munn, J., & Pier, J. 2006, SDSS Astrometric
     Pipeline documentation, available on request from Dr. Jeff Pier

Hirschel, F. & Banks, J. 1782, Phil. Trans., 72, 82

Hirt, C. 2006, A&A, 459, 283

Hudson, C. J. 1929, Pub. Allegheny Obs., 6, 1

Ivezic, Z., *et al.* 2008, astro-ph/0805.2366

Kaiser, N., Tonry, J. L., & Luppino, G. A. 2000, PASP, 112, 768.

Kaiser, N., *et al.* 2002, Proc. of SPIE, 4836, 154

Kaplan, G. H. & Bangert, J. A. 2006, IAUJD, 16, #3

Kitchin, C. R. 2003, Astrophysical Techniques (Philadelphia: IOP)

Knuteson, R. O., *et al.* 2004, J. Atm. & Ocean. Tech., 21, 1765

Kolmogorov, A. 1941, Dokl. Akad. Nauk SSSR, 30, 301

Kovalevsky, J. & Seidelmann, P. K. 2004, Fundamentals of Astrometry (Cambridge:
    Cambridge University Press)

Küstner, F. 1888, Baob.-Erg. Sternwerte Berlin, 3, I-59

Lambert, W. D., Schlesinger, F., & Brown, E. W. 1931, Bull. NRC, 78, 269

Land, G. 1944, AJ, 51, 25

Land, G. 1954, AJ, 59, 19

Lee, X. 1997, J. Atm. Sci., 54, 2574

Lieske, J. H., Lederle, T., Fricke, W., & Morando, B. 1977, A&A, 58, 1

Lu, C. & Koch, S. E. 2008, J. Atm. Sci., 65, 2676

Magnier, E. A., Liu, M., Monet, D. G., & Chambers, C. 2007, IAUP, 3, 553

Martner, B. E., *et al.* 1993, BAMS, 74, 599

McGraw, J. T., *et al.* 2005, AMOS conf. proc., 382

Mink, D. J. 1997, ASPC, 125, 249

Nappo, C. J. 2002, An introduction to atmospheric gravity waves (San Diego: Elsevier
    Science)

Newsom, R. K. & Banta, R. M. 2003, J. Atm. Sci., 60, 16

Perrine, C. D. 1913, AJ, 28, 31

Pier, J. R., Munn, J. A., Hindsley, R. B., Hennessy, G. S., Kent, S. M., Lupton, R. H., &
Ivezic, Z, 2003, AJ, 125, 1559

Poulos, G. S., *et al.* 2002, BAMS, 83, 555

Roddier, F. 1999, Adaptive Optics in Astronomy (Cambridge: Cambridge University
Press)

Roggeman, M. C. & Welsh, B. 1996, Imaging Through Turbulence (Boca Raton: CRC
Press)

Schlesinger, F. 1899, PASP, 11, 230

Schlesinger, F. 1905, AJ, 24, 183

Schlesinger, F. 1916, Pub. Allegheny Obs., 3, 1

Schlesinger, F. & Blair, G. B. 1906, PASP, 18, 46

Schlesinger, F. & Hudson, C. J. 1916, Pub. Allegheny Obs. 3, 55

Scorer, R. S. 1997, Dynamics of Meteorology and Climate (Chichester: Praxis)

The SDSS Project Book. 1999, The Astrophysical Research Consortium,
http://www.astro.princeton.edu/PBOOK/welcome.htm

Smith, W. L., Feltz, W. F., Knuteson, R. O., Revercomb, H. E., Woolf, H. M., & Howell,
H. B. 1999, J. Atm. & Ocean. Tech., 16, 323

Sozzetti, A. 2005, PASP, 117, 1021

Stetson, P. 1987, PASP, 99, 191

Stone, R. C., Monet, D. G., Monet, A. K. B., Walker, R. L., Ables, H. D., Bird, A. R., &
Harris, F. H. 1996, AJ, 111, 4

Stull, R. B. 1988, An introduction to atmospheric boundary layer meteorology

    (Dordrecht: Kluwer Academic Publishers)

Sugawa, C. 1956, PASJ, 8, 27

Sugawa, C. 1958, PASJ, 10, 79

Sugawa, C. 1960, PASJ, 12, 42

Tatarski, V. I. 1961, Wave propagation in a turbulent medium (New York: Dover,

    translated by R. A. Silverman)

Taylor, G. I., 1938, Proc. Royal Soc. London, 164, 476

Thompson, L. A. & Gardner, C. S. 1987, Nature, 328, 229

Tyson, J. A. 2002, Proc. of SPIE, 4836, 10

Vangeyte, B., Manfroid, J., & Surdej, J. 2002, A&A, 388, 712

Watts, C. B. 1960, in Telescopes (Chicago: The Univeristy of Chicago Press, G.B.

    Kuiper & B. M. Middlehurst eds.)

Wells, N. 1997, The atmosphere and ocean: a physical introduction (Chichester: John

    Wiley & Sons)

York, D. G., *et al.* 2000, AJ, 120, 1579

Zacharias, N., Monet, D. G., Levine, S. E., Urban, S. E., Gaume, R., & Wycoff, G. L.

    2004, BAAS, 36, 1418

Zacharias, N., *et al.* 2000, AJ, 120, 2131

Zhang, S. D. & Yi, F. 2005, Ann. Geophys., 23, 665

Zimmer, P. C., McGraw, J., Gimmestad, G., Roberts, D., Smith, J., & Fitch, J. 2009,

    AAS 41, 433